DOCUMENT RESUME

ED 363 089 FL 019 686

AUTHOR Kokkota, V. A.
TITLE Lingvo-didakticheskoe testirovanie (Language Testing).
REPORT NO ISBN-5-06-000219-5
PUB DATE 89
NOTE 130p.
AVAILABLE FROM Vysshaya Schola Publishing House, Neglinnaya St. 29/14, Moscow 101430 Russia.
PUB TYPE Books (010) -- Reports - Descriptive (141)
LANGUAGE Russian

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Communicative Competence (Languages); Foreign Countries; Language Skills; *Language Tests; Linguistic Competence; Modern Language Curriculum; Scoring; Second Language Learning; Test Construction; *Testing; Test Interpretation; *Test Theory
IDENTIFIERS USSR

ABSTRACT

This book contrasts non-Soviet approaches to language testing and provides definitions from four Soviet language test experts. The role of foreign language teaching, the function of tests, and theoretical problems are discussed, with considerable focus on communicative competence. The book discusses test standardization and classification and treats statistical and arithmetical handling of data. It also contrasts brief histories of language testing abroad and in the Soviet Union. Specific attention is focused on theory and general methodology of test development and interpretation of results. The following issues are highlighted: testing for lexical and grammatical knowledge or competencies; testing for reading or aural comprehension; and pragmatic tests (dictations, aural and visual cloze tests, speaking and oral production tests). A bibliography with 170 citations is included. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 363 089

БП

В. А. Коккота

ЛИНГВО-ДИДАКТИЧЕСКОЕ ТЕСТИРОВАНИЕ

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

V. Kokkota

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

FL 019686

Библиотека
преподавателя

В. А. Коккота

ЛИНГВО-ДИДАКТИЧЕСКОЕ ТЕСТИРОВАНИЕ

Москва
«Высшая школа»
1989

ББК 81.2 Англ
К 55

Рецензенты:
кафедра иностранных языков Московского экономико-статистического института (зав. кафедрой проф. В.В. Морозенко); д-р филол. наук, проф. Ю.А. Тулдава (Тартуский государственный университет)

Рекомендовано к изданию Государственным комитетом СССР по народному образованию

Коккота В.А.
К 55 Лингводидактическое тестирование: Науч.-теорет. пособие. – М.: Высш. шк., 1989. – 127 с. – (Б-ка преподавателя).
ISBN 5-06-000219-5

Основное внимание в пособии уделено анализу отечественных и зарубежных тестов для иллюстрации практических методов их составления и использования, а также обработки результатов. Рассмотрены проблемы тестирования во всех основных видах речевой деятельности и аспектах языка, возможности тестирования для само- и взаимоконтроля обучаемых.
Пособие имеет справочный характер, позволяющий преподавателям качественно и экономно составлять и использовать языковые тесты.

К $\frac{4602020102\ (4309000000) - 281}{001\ (01) - 89}$ 286–89 ББК 81.2 Англ

ISBN 5-06-000219-5 © Издательство «Высшая школа», 1989

1. ПРЕДИСЛОВИЕ

В свете перестройки народного образования в нашей стране перед высшей школой стоит задача обеспечить подготовку специалистов на современном уровне, обладающих основательными теоретическими знаниями и практическими навыками.

Указанное положение налагает большую ответственность и на преподавателей иностранных языков. Растущие внешнеэкономические связи СССР требуют привлечения специалистов со знанием одного или нескольких иностранных языков, причем с лучшим, чем раньше, владением устной речью. В связи с этим возрастает роль коммуникативно-ориентированной методики, интенсивных и ускоренных методов обучения иностранным языкам, а также самостоятельной работы обучаемых, в том числе с использованием ЭВМ и видеотехники.

Внедрение коммуникативно-ориентированной методики обучения иностранным языкам влечет за собой новые задачи по контролю обученности, в том числе по тестовому контролю. С одной стороны, растет роль прагматических и коммуникативных тестов, с другой – требуется разработка новых тестов для проведения научных экспериментов, чтобы решить ряд методических проблем.

О возрастающем значении тестирования в преподавании иностранных языков говорит следующий факт: в библиографии Д. Ланге и Р. Клиффорда [128] приведено около 1500 публикаций по иноязычному тестированию, которые появились в период 1966–1979 гг. В нашей стране прослеживается аналогичная тенденция, только масштабы намного меньше. Последнее можно объяснить как историческими причинами, так и тем, что образование в нашей стране не является тесто-ориентированным. Однако, вряд ли стоит недооценивать такие выгодные качества тестов, как экономичность, надежность и объективность подсчета их результатов и др.

В настоящем пособии сделана попытка по-новому осветить как теоретические вопросы, так и практические проблемы тестирования. Теоретические вопросы связаны с коммуникативной теорией языка и с компонентной структурой коммуникативной компетенции учащихся; практические проблемы связаны с разделением тестов на языковые и речевые (их общее название – лингводидактические тесты), с использованием шкал стандартизованности, контекстуальности и коммуникативности лингводидактических тестов, а также с некоторыми чисто тесто-

техническими вопросами, как, например, с экономичностью, качеством и субъективностью оценки ответов на тестовые задания. Все вопросы теории и практики рассмотрены с точки зрения системы: тест, тестируемый, тестолог и условия (ситуация) тестирования. Более половины пособия посвящено вопросам планирования, составления, проведения, обработки и интерпретации результатов лингводидактических тестов. Хотя основное внимание уделено тестам по лексике, грамматике, чтению и аудированию, описаны также проблемы составления тестов устной речи, тестов типа клоуз (дополнения) и вопросы тестирования с помощью ЭВМ.

Из-за малого объема пособия исторический обзор развития тестирования очень краток, эта проблема более подробно освещена в работах двух наших ведущих тестологов: И.А. Цатуровой [77] и И.А. Рапопорта и др. [62]. Многие данные приведены в таблицах и рисунках, дано краткое изложение методики обработки результатов теста и небольшое количество примеров, иллюстрирующих разные типы заданий и тестов.

В настоящем пособии использован опыт отечественной и зарубежной тестологии, в частности, опыт автора и руководимого им коллектива в Таллинском политехническом институте, где завершается составление и опробование вузовской системы стандартизованных тестов, составлены локально стандартизованные тесты по устной речи, аудированию и некоторых вербальных способностей студентов.

По охвату и уровню презентации материала пособие полезно как для специалистов в области тестирования, так и для широкого круга – преподавателей вузов, учителей школ, студентов-филологов, преподавателей языковых курсов.

Автор

2. МЕСТО ТЕСТИРОВАНИЯ В ПРЕПОДАВАНИИ ИНОСТРАННЫХ ЯЗЫКОВ

2.1. ОСНОВНЫЕ ПОНЯТИЯ В ОБЛАСТИ ЛИНГВОДИДАКТИЧЕСКОГО ТЕСТИРОВАНИЯ

Такие понятия, как тестирование, тест, тестируемый и др., повторяются многократно во всех разделах настоящего пособия, однако в литературе, как будет показано ниже, их содержание не совсем однозначно. Поэтому для удобства чтения англоязычной литературы некоторые понятия приведены ниже с соответствиями из английском языке.

Тестирование (testing) как термин встречается в литературе в узком смысле в значении использования и проведения теста и в широком смысле как совокупность процедурных этапов планирования, составления и опробования тестов, обработки и интерпретации их результатов. В данном пособии термин «тестирование» используется только в широком смысле.

Приведенный перечень этапов тестирования не является полным, и полная временнáя последовательность этапов тестирования приведена на рис. 1, причем первая стадия разработки предтеста может повториться еще раз в случае его двукратного опробования. Подробно эти этапы тестирования будут рассмотрены в разделе 5.

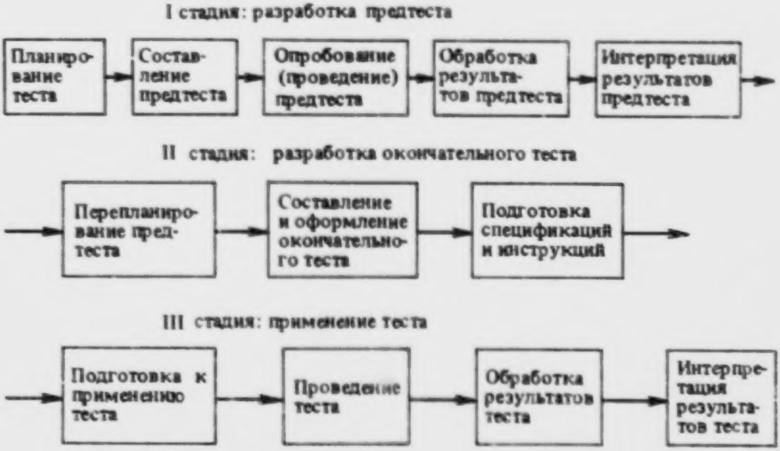

Рис. 1. Последовательность этапов тестирования.

Лингводидактическое тестирование (language testing) занимается разработкой и использованием языковых и речевых тестов и, с одной стороны, является областью методики преподавания иностранных языков, а с другой стороны, относится к педагогическому (предметному) тестированию как часть к общей тестологии. В общую тестологию входят также психологические, профессиональные и другие тесты.

Тестология (testing) – это теория и практика тестирования, а **тестолог** – это специалист, занимающийся теорией и практикой тестирования. **Тестирующим** (tester) называют человека (преподавателя), который проводит тест среди **тестируемых или участников теста** (test-taker, testee), т.е. среди людей (учащихся, студентов и т.д.), которые выполнили или будут выполнять тест.

На рис. 2 приведены основные элементы комплекса разработки и применения теста. Исходные данные для планирования теста тестолог получает из учебного процесса, применяемой методики и средств обучения (учебник, ТСО и т.п.). К основным элементам указанного комплекса относятся, кроме тестологов, еще преподаватели и администраторы (например, заведующие кафедрой), которые применяют готовый тест, сам тест со своими данными и два контингента тестируемых, причем меньшая часть (репрезентативная выборка) может впоследствии войти в основной контингент.

Рис. 2. Основные элементы комплекса разработки и применения теста.

Слово **тест** в английском языке обозначает «испытание, проба, эксперимент, проверка». При такой полисемии неудивительны терминологические неточности, и иногда сравнительно простые контрольные работы и контрольно-тренировочные упражнения также называются тестами, что не совсем оправдано, поскольку преподавательские тесты тоже должны быть составлены в соответствии с определенными правилами и требованиями.

Как указывает Ю.З. Гильбух [12, с. 31], термин «тест» еще не имеет достаточно точного определения в психологии. Это очевидно из весьма

общих определений психологического теста: «Систематизированная процедура для наблюдения и описания поведения индивида с помощью шкалы или понятийной системы» [101, с. 6]; «объективное и стандартизованное измерение выборки поведения» [1, с. 31] и др.

В зарубежной литературе по лингводидактическим тестам практически не обсуждается вопрос о сущности теста, поскольку контроль в зарубежной методике преподавания иностранных языков (ПИЯ) в основном осуществляется путем либо тестов, либо экзаменов. В отечественной литературе по лингводидактическим тестам имеется четыре определения, которые заслуживают внимания (приводятся в хронологическом порядке появления в печати):

1. С.И. Воскерчьян [8] считает, что «тест – это кратковременное, технически просто обставленное испытание, проводимое в равных для всех условиях и имеющее вид такого задания, решение которого поддается количественному учету и служит показателем степени развития к данному моменту известной функции у данного испытуемого».

2. И.А. Цатурова в своей диссертации [79] приводит следующее определение теста: «тест – это централизованный, формализованный письменный контроль, который дает возможность определить основные параметры качественного и количественного состояния обученности учащихся на каждом этапе обучения».

3. Э.А. Штульман [82], считая тест методом эксперимента, определяет тест так: «методический тест – это контрольное задание (задания), проводимое в равных для всех обучающихся условиях, длительность и характер которого строго соотнесены с объективными факторами и результаты выполнения которого поддаются количественной оценке, являясь показателем определенных итогов учебного процесса к моменту данного теста».

4. И.А. Рапопорт и др. [62] предлагают следующее определение теста: «тест – это самым тщательным образом подготовленная в соответствии с определенными разработанными правилами, прошедшая предварительную экспериментальную проверку и специальную процедуру для ее улучшения, имеющая достаточные характеристики своей эффективности совокупность вопросов и заданий, предъявляемых испытуемому с целью квалиметрического выявления социальных, психических и психофизиологических характеристик его личности, отличающаяся формализацией ответов испытуемых, выделением в них части, несущей наибольшую информационную нагрузку, что ускоряет, облегчает и объективизирует их последующий анализ, обработку и интерпретацию».

Все приведенные выше определения теста имеют тот общий недостаток, что не допускают возможности составления относительно субъективных речевых тестов, особенно таких, где затруднена или невозможна формализация ответов. Кроме того, определение С.И. Воскерчьян слишком «психологическое», оно не учитывает специфику лингводидактических тестов и считает тестом только кратковременное испытание, в то время как стандартизованные тесты могут длиться 1–2 часа. В опре-

делении И.А. Цатуровой тестом считается только письменный контроль, а сейчас имеются и разрабатываются тесты устной речи (коммуникативной компетенции). Как и в определении Э.А. Штульмана, здесь опущено упоминание о подготовительной стадии составления теста (планирование, составление, опробование), что имеется в определении И.А. Рапопорта и др. Укажем здесь, что разница между контрольной работой программированного типа и аналогичным тестом, называемым ниже преподавательским (учительским) тестом, заключается как раз в том, что задания теста предварительно опробованы с целью выявления их качества и повышения эффективности теста.

Как следует из приведенного анализа имеющихся определений теста, весьма противоречивые характеристики языковых и речевых, а также преподавательских и стандартизованных тестов не позволяют сформулировать достаточно конкретное определение лингводидактического теста. Ниже под **лингводидактическим тестом** понимается подготовленный в соответствии с определенными требованиями комплекс заданий, прошедший предварительное опробование с целью определения его показателей качества и позволяющий выявить у тестируемых степень их языковой (лингвистической) и/или речевой (коммуникативной) компетенции, и результаты которого поддаются определенной оценке по заранее установленным критериям.

Уточним данное определение теста. В него введено понятие комплекса заданий, поскольку характер заданий даже языковых тестов носит разнообразный характер. Например, в грамматических тестах требуется также определенное знание лексики, на основе которой эти задания составлены. Требования к подготовке теста относятся к уточнению целей и объектов тестирования, к анализу охватываемого языкового материала, к разработке процедур и инструкций к проведению теста и т.п. Остальные характеристики лингводидактических тестов будут подробнее рассмотрены ниже.

Крупные тесты с многими (более 40–50) заданиями называются **комплексными тестами или тестовыми батареями** (test battery) и состоят из **частей** (part) и **субтестов** (subtest). Субтест состоит из однотипных заданий и составляется на один конкретный **объект тестирования** (test task). Под последним понимается соответствующий цели тестирования компонент иноязычных знаний, навыков и умений. Часть теста составляется на укрупненный объект тестирования, например, на чтение, аудирование и т.п.

Тестовым заданием или **заданием теста** (item) называют ту минимальную составляющую единицу теста, которая предполагает определенную вербальную или невербальную реакцию (answer, response) тестируемого.

Каждое тестовое задание создает для тестируемого некоторую лингвистическую и экстралингвистическую ситуацию, называемую тестовой ситуацией, точнее – ситуацией тестового задания. Тестовая ситуация представлена стимулом задания (stimulus), который может быть вербаль-

ным или невербальным, т.е. наглядным. Вербальный стимул в тестовом задании называют основой (stem), или текстом задания. Ответ (response) тестируемого также может быть вербальным или невербальным. Вербальный ответ может быть выборочным или, как его называют иногда, – избирательным (multiple-choice response) или свободно-конструируемым (free response). В первом случае предлагается выбрать ответ среди вариантов (answers, choices), один из которых является правильным (correct answer), другой/другие – отвлекающими, или дистракторами (distracters).

Планирование состава теста, выбор объектов тестирования и типа заданий и т.п. называют техникой тестирования (testing technique).

Остальные определения используемых терминов будут приведены по мере их появления в тексте пособия.

2.2. РОЛЬ ТЕСТИРОВАНИЯ В ПРЕПОДАВАНИИ ИНОСТРАННЫХ ЯЗЫКОВ

Начиная с 70-х годов, в отечественной практике преподавания иностранных языков роль тестирования постоянно растет. Свидетельство этому – рост количества публикаций, защит диссертаций, проводимых семинаров по тестированию и т.п., а также больщой интерес преподавателей и учителей к тестированию и большее применение тестов в их повседневной учебной работе.

Кроме преподавательских и обучающих тестов большое значение имеют лингводидактические тесты, особенно для решения таких ключевых проблем методики ПИЯ, как, например, проблемы познания процесса усвоения ИЯ, понимание сущности коммуникативной компетенции, проблемы оптимизации методики преподавания ИЯ данному контингенту учащихся в конкретных условиях обучения и др.

Основная роль тестирования в преподавании ИЯ заключается в обратной связи (в широком смысле) и в контроле (в узком смысле), и, чем теснее взаимосвязаны тестирование и учебный процесс, тем эффективнее учебный процесс и качественнее тестирование. На рис. 3 приведены некоторые этапы тестирования с двумя ветвями обратной связи. Из них нижние ветви $ВБ_2$ и $ВБ_3$ – традиционные, которые обычно имеют место и при нетестовой методике контроля, а верхняя ветвь ВГ свойственна в основном только тестам, т.е. результаты теста используются для улучшения или проверки качества следующего или данного варианта этого теста. Под качеством здесь понимается валидность и надежность теста, а также приемлемая трудность и дифференцирующая способность заданий теста. Эти показатели тестов будут подробно рассмотрены ниже.

Схема связей на рис. 3 является неполной в том смысле, что подобных циклов по крайней мере два – один для предварительного теста (предтеста) и второй для окончательного теста. Кроме того, чтобы показать более полную картину обратной связи, на рис. 3 изображен некий идеализированный тест, позволяющий получить информацию о достижениях

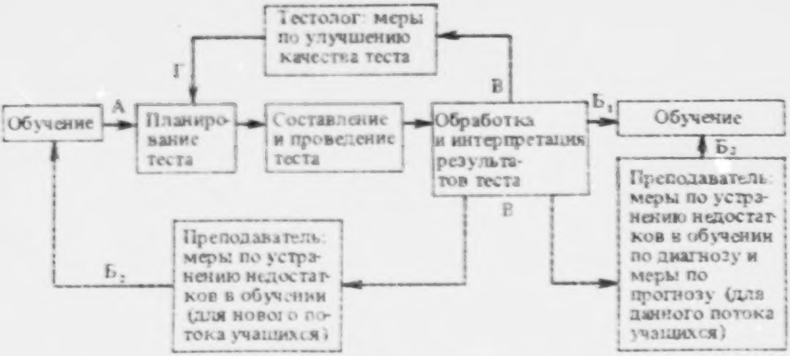

Рис. 3. Схема обратной связи в процессе тестирования.

обучения, о диагнозе пробелов и о прогнозе успешности обучения. Реальный тест позволяет получить в основном один из указанных видов информации.

Вопрос о роли тестирования в практике преподавания ИЯ во многом сводится к раскрытию функции тестов (см. п. 2.3), преимуществ и недостатков тестов (см. п. 3.1.2), а также к соотношению тестовых и нетестовых форм контроля. Указанное соотношение в основном зависит от применяемой методики обучения ИЯ, а также от наличия подходящих тестов, возможностей их разработки и от других обстоятельств. Очевидно, что, чем выше административный и организационный уровень контроля и соответственно больше охват обучаемых и учебного материала, тем выше потребность именно в тестовой методике контроля. Забегая вперед, можно сказать, что тем выше должна быть степень стандартизации теста.

2.3. ФУНКЦИИ ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

Все функции ЛДТ[1] можно разбить на следующие общие группы: 1) общедидактические функции обучения (собственно обучающая, повторительно-закрепляющая и развивающая) и функции воспитания (познавательная, дисциплинирующая и др.); 2) информационные функции результатов ЛДТ (контрольная, исследовательская и учебно-информационная). Эти функции частично перекрываются, например, важнейшая функция ЛДТ – контрольная – относится традиционно и к группе

[1] Функции тестов подробно освещены в статьях И.А. Цатуровой [78] и Н.С. Шебеко [80], некоторые функции ЛДТ рассмотрены нами в разделе классификации ЛДТ.

общедидактических функций. Кроме того, среди исследовательских функций ЛДТ имеются функции развития памяти, внимания и личностных качеств (добросовестности, целеустремленности и т.п.), последние из которых относятся и к воспитательной функции ЛДТ.

На рис. 4 сделана попытка показать вышеуказанные функции ЛДТ в связи с различными этапами процесса тестирования. Ради ясности на рис. 4 пришлось опустить ряд связей и повторяющихся на разных этапах тестирования функций ЛДТ.

Общедидактические функции обучения проявляются уже в ходе подготовки к тесту. Обучаемые систематизируют свои знания, ликви-

Рис. 4. Основные функции лингводидактических тестов.

дируют пробелы, обнаруженные в ходе само- и взаимоконтроля и т.п. В ходе выполнения теста тестируемые выполняют разные речевые действия (деятельностная функция), при этом происходит определенное повторение и закрепление учебного материала. Активизируется мыслительная деятельность, развивается память, т.е. достигается определенный обучающий эффект. Этот эффект тем больше, чем более контекстуален и коммуникативен тест, чем быстрее происходит обратная связь в системе тест – тестируемый, например, при использовании ЭВМ.

Как указывает С.К. Фоломкина [76], можно составить и специальные обучающие тесты. Таковыми могут быть, например, тесты типа клоуз, диктанты и др. При анализе результатов теста также достигается обучающий эффект теста (операции типа: проверка по ключу ошибочных ответов). В будущем, вероятно, роль обучающей функции ЛДТ будет расти (более коммуникативные тесты, применение ЭВМ и т.п.).

ЛДТ выполняют воспитательную функцию также на этапе подготовки к тесту и в ходе работы над заданиями теста через преодоление различных психолингвистических, мыслительных и других трудностей, т.е. путем воспитания силы воли, целеустремленности, добросовестности и других качеств личности, а также путем изменения отношения к учебе, к иностранному языку. Эмоциональное состояние тестируемых изменяется также благодаря стимулирующей роли оценки (балла) теста (при анализе результатов теста). Заметное воспитательное воздействие оказывает применение ЭВМ для проведения теста.

Основной функцией тестирования является функция контроля, которая на рис. 4 отделена от учебно-организационной функции по тому принципу, что при проведении тестов для контроля успеваемости могут и не приниматься организационные решения (имеет место только констатация успеваемости). В случае тестов, составленных для диагноза, прогноза и т.п. успеваемости, указанные решения непременно принимаются, например, для более обоснованного распределения обучаемых в учебные группы или для корректировки рабочих программ.

Благодаря дополнительной обработке их результатов ЛДТ могут выполнить специфическую функцию выявления качества теста (надежности и валидности). Если ЛДТ составлялись для использования в виде срезов для методических экспериментов, тестов типа клоуз для оценки трудности («читабельности») учебных текстов или тестов способностей, то такие тесты выполняют специфические исследовательские функции.

Для максимально эффективной реализации вышеуказанных функций ЛДТ следует соблюдать следующие основные принципы:

1) составление или применение только таких ЛДТ, которые достаточно качественны, достаточно надежны, валидны с точки зрения целей и задач тестирования и учебно-воспитательной работы; 2) соблюдение обычных процедур разработки, проведения и обработки результатов ЛДТ, изложенных в литературе и ниже в настоящем пособии; 3) соблюдение оптимального соотношения тестовых и нетестовых форм контроля, т.е. использование ЛДТ только тогда, когда это оправдано; 4) со-

блюдение общедидактических принципов контроля – регулярности, систематичности, активности, сознательности и др. при организации тестового контроля.

3. ТЕОРЕТИЧЕСКИЕ ПРОБЛЕМЫ ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

3.1. ОБЩАЯ ХАРАКТЕРИСТИКА ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

3.1.1. Сравнение ЛДТ с психологическими и педагогическими тестами

Первые языковые тесты появились в середине 20-х годов нашего столетия. Их организация и техника были заимствованы из психологических и педагогических тестов (ППТ). Лингводидактические тесты и ППТ имеют следующие общие свойства:

1. Во всех трех видах тестов взаимодействуют элементы следующих трех систем: 1) тест, средства и условия тестирования; 2) составители теста (тестологи), пользователи тестом (например, преподаватели), администраторы (например, заведующие кафедрами) и общество; 3) процессы составления и проведения (выполнения) теста, обработки и интерпретации его результатов. Указанное взаимодействие обусловливает определенную субъективность результатов тестов, что имеет место даже в случае так называемых объективных тестов.

2. Эти разновидности тестов измеряют определенные характеристики человеческой деятельности (поведения, мышления, речи и т.п.) и делают это только в ходе человеческой деятельности на основе определенной выборки из разных возможных видов деятельности.

Из указанного выборочного характера тестов следует относительная неполнота результатов измерений с помощью ЛДТ и ППТ и ограниченные возможности их обобщения. Кроме того, поскольку измеряется не сама характеристика, а ее проявление в деятельности, то результаты любых измерений с помощью этих тестов имеют относительно косвенный характер.

3. Измерение указанных характеристик с помощью ППТ и ЛДТ производится в сложной системе «человек» с многими взаимосвязанными и поэтому взаимно влияющими характеристиками, что обусловливает значительную методологическую погрешность измерений и вероятностно-статистический характер результатов тестов.

Учет указанных свойств относительной субъективности и неполноты результатов этих тестов, их косвенный и приближенный (вероятностный) характер важен при планировании и составлении тестов и особенно при интерпретации их результатов.

4. Все эти разновидности тестов находятся на общей шкале (континууме) жизненного опыта тестируемых, который влияет на качество и результат деятельности тестируемого в ходе выполнения теста. На одном конце указанной шкалы находятся ЛДТ учебных достижений и педагогические тесты, а на другом конце – невербальные и проективные психологические тесты. Между указанными тестами находятся ЛДТ общего владения ИЯ, тесты лингвистических способностей, тесты интеллекта и другие.

ЛДТ имеют следующие отличительные свойства:

1) если в случае психологических тестов основное внимание обращается на то, как провести тестирование, то в случае ЛДТ основное внимание обращается на то, что тестировать. Объектом тестирования в случае ЛДТ является коммуникативная компетенция – знания, навыки и умения тестируемых в их речевых действиях, т.е. ЛДТ имеют непременно вербальный характер (либо задание, либо ответ является вербальным);

2) ЛДТ имеют характерные для педагогических тестов признаки измерения обученности, развития и прогресса в учебной деятельности, а психологические тесты измеряют скорее обучаемость и связанные с ней параметры, причем в случае ЛДТ есть возможность построения шкалы абсолютной оценки с помощью понятия образованного носителя рассматриваемого ИЯ;

3) задания ППТ независимы друг от друга и в основном измеряют только одно умение, т.е. они узконаправлены, чего нельзя сказать о ЛДТ, особенно при высокой степени коммуникативности.

3.1.2. Преимущества и недостатки лингводидактических тестов

Ниже критически рассмотрены некоторые преимущества и недостатки, присущие ЛДТ. (Недостатки ЛДТ, которые возникают при их составлении, будут рассмотрены в соответствующих разделах пособия.)

Иногда ЛДТ наделяют такими преимуществами, как возможность проконтролировать одновременно всю группу учащихся, применяя при этом одинаковый для всех учебный материал и одинаковые для всех условия (с точки зрения времени и места). Указанные качества, однако, проявляются только в случае сравнения ЛДТ с индивидуальным устным опросом, который не является самым эффективным видом нетестового контроля. Например, проверку лексики, грамматики, чтения можно организовать как с помощью ЛДТ, так и с помощью письменной контрольной работы с нетестовыми заданиями. Последняя имеет все вышеуказанные качества ЛДТ.

Бесспорными преимуществами ЛДТ перед нетестовыми формами контроля являются: 1) более высокое качество заданий благодаря их предварительной проверке в предтесте; 2) заметно бо́льшая надежность ЛДТ благодаря выборочному способу ответов, что обеспечивает объективность подсчета результатов теста, а также за счет возможности при-

влечения большего количества заданий за одинаковое время и по ряду других причин; 3) относительная объективность всего языкового теста и объективность подсчета его результатов; 4) экономия аудиторного времени за счет компактной письменной формы заданий языкового теста, что дает возможности включения в тест большого количества заданий; 5) экономичность проверки ответных листов ЛДТ, т.е. возможность проверки ответов тестируемых на задания с выборочными ответами с помощью специальных трафаретов и ЭВМ, что проявляется, однако, лишь при достаточно массовом и многократном использовании ЛДТ; 6) возможность стандартизации процедур проведения тестов и применения статистического анализа результатов языковых тестов; 7) более положительная мотивация и отношение тестируемых к тестовым формам контроля, так как исключена роль субъективности оценки, момент невезения с вопросами-заданиями и т.п.; 8) сравнительная легкость перевода языковых и части речевых тестов на ЭВМ, что может дать этим тестам дополнительные преимущества; 9) удобство массового использования языковых тестов; 10) лучшая воспроизводимость и репрезентативность результатов ЛДТ благодаря однообразной (стандартной) процедуре тестирования.

Недостатки и отрицательные стороны применения ЛДТ связаны с некоторыми предубеждениями и недопониманием свойств тестов, среди которых отметим следующие. Считается, что тесты вообще – это нечто сомнительное, некачественное и загадочное, а собственно ЛДТ – это обычное контрольное задание особой формы и улучшенного качества, пригодность которого для данных условий зависит от его составителей и пользователей. Иногда ЛДТ присваивают качества, которыми они не обладают, например, обычный перевод или сочинение также называют тестом. Часто считается, что тест легче, чем аналогичное нетестовое задание из-за той возможной подсказки, которая имеет место в случае выборочных ответов на тестовое задание.

Иногда, наоборот, имеет место абсолютизация и идеализация значения и положительных качеств ЛДТ. Отметим, что ЛДТ имеет целый ряд недостатков, с которыми следует считаться при целенаправленном сочетании тестовых и нетестовых форм контроля.

Общим недостатком языковых и речевых тестов является их уменьшенная внешняя валидность. В случае языковых тестов это вытекает из искусственности изолированных от большого контекста заданий, а также из-за выборочных ответов. Если рассматривать тесты по устной речи, то здесь внешняя валидность высокая, но не максимальная с точки зрения коммуникативной теории языка. Дело в том, что практически даже в идеальном случае не удается обеспечить аутентичности во всех элементах системы – тестовая ситуация (тема, коммуникативная задача и т.п.), характер используемого языка, состояние тестируемого и рейтера, оценивающего ответы тестируемого. Особенно это касается рейтера, поскольку в реальной коммуникации партнеры сотрудничают, помогают друг другу в обмене информацией, а в случае теста рейтеру следует еще

некоторым образом произвести оценку развитости тех или иных навыков и умений. Положение аналогично дилемме квантовой физики – влияет ли измерительный прибор на результат измерения. Похоже, что в случае ЛДТ – влияет.

Главным недостатком речевых тестов без выборочных ответов является субъективность оценки ответов тестируемых, что снижает надежность тестов этого вида. Ниже будет показано, что с помощью регламентированных процедур проведения и оценки результатов указанных тестов от этого недостатка удается практически избавиться.

Следующие недостатки языковых тестов связаны с использованием выборочных ответов:

1. Применяемая обычно методика представления нескольких неправильных вариантов ответов и одного правильного не соответствует типичной для реальной коммуникации стратегии. В общении исходят либо из наиболее подходящего для данной ситауции выбора (особенно в случае родного языка) или из наличных вариантов для поддержания общения (в случае ИЯ).

2. Угадывание правильного ответа как по случайным, так и по формальным или другим, не относящимся к владению языком, признакам (если ответ сразу не найден). Случайное угадывание обычно ограничивается стремлением тестируемого найти правильный ответ путем построения, подтверждения или отвергания гипотез о пригодности того или иного варианта ответа. Доказательством сказанного являются 10–15 % правильных ответов у некоторой части тестируемых при величине случайного угадывания 25 % при 4-х вариантах ответов. Задания языкового теста можно составить таким образом, чтобы угадывание по внешним и другим признакам было минимальным.

3. Недостаточная активность мыслительных процессов из-за наличия готовых ответов и из-за общей ориентации на результат теста. Указанную активность можно несколько повысить с помощью подбора соответствующих характеру объектов тестирования типов и форм заданий, а также регулированием степени трудности заданий.

4. Возможность непроизвольного запоминания неправильных ответов. Такая опасность существует при слишком трудных заданиях, когда тестируемый не может построить гипотезы о правильном варианте ответа. Поэтому на начальных этапах обучения предлагается больше использовать невербальные (наглядные) формы ответов.

5. Тестовая ситуация в задании недостаточно полна, информативна и коммуникативна из-за сжатости контекста в языковых тестах (1–2 предложения). С помощью специальных мер для улучшения смысловой и экстралингвистической стороны задания этот недостаток обычно удается в значительной степени исправить.

6. Неправильные варианты ответов могут содержать избыточную и отвлекающую информацию, препятствующую определению правильного ответа. В этом отношении составителю теста следует быть весьма осмотрительным, чтобы соблюдать так называемое семантическое равновесие среди вариантов ответов тестового задания.

3.2. ЛИНГВИСТИЧЕСКАЯ И КОММУНИКАТИВНАЯ КОМПЕТЕНЦИЯ

Лингвистическая компетенция. Термин «компетенция» происходит от латинского competere *быть способным к чему-либо.* Лингвистическая компетенция (ЛК) по Н. Хомскому означает способность понимать/продуцировать неограниченное число правильных в языковом отношении предложений с помощью усвоенных языковых знаков и правил их соединения (по статье Н.И. Гез [11]). Р. Валетт [166] связывает ЛК с объемом (breadth) и точностью владения лингвистическими элементами (фонологией, грамматикой и лексикой). М. Финоккиаро [108] определяет ЛК как способность понимать и продуцировать изученные или аналогичные им выражения (utterances), а также потенциальную способность понимать новые, неизученные выражения. Р. Шульц [149] дает определение ЛК как способности воссоздавать (recall) и манипулировать дискретными лингвистическими фактами о фонологии, морфологии и синтаксисе изучаемого языка.

В методической литературе имеются и другие определения ЛК, однако все они сравнительно близки друг к другу, и, главное, ЛК рассматривается как составляющий компонент коммуникативной компетенции (КК). Выявление ЛК у тестируемых не представляет значительных теоретических проблем, поскольку на всех стадиях тестирования используются известные психометрические методы и приемы. В таблице 1 приведены теоретически возможные виды языковых тестов для определения компонентов лингвистической компетенции. Согласно этой таблице можно составить тесты по грамматике для чтения, по лексике для говорения и т.д. Компоненты лингвистической компетенции будут обсуждены ниже при рассмотрении тестирования навыков и умений чтения, аудирования и т.д.

Таблица 1

Теоретически возможные виды языковых тестов лингвистической компетенции

Аспекты языка	Рецептивные виды речевой деятельности		Продуктивные виды речевой деятельности	
	чтение	аудирование	письмо	говорение
Фонология (орфография)				
Грамматика				
Лексика				
Экстралингвистический (скорость, темп ...)				

Понятие коммуникации. В определениях КК гораздо меньше единогласия, и это вытекает из того факта, что в определении сущности самого процесса коммуникации нет должной ясности. Для правильной организации тестирования полезно сначала рассмотреть общее определение и те характеристики коммуникации, которые связаны с сущностью языка и речевой деятельности.

В теории информации коммуникацией считают деятельность, в результате которой происходит определенное уменьшение энтропии, т.е. определенное увеличение упорядоченности информации. Нас интересует такая упорядоченность информации, которая происходит в процессе общения между людьми. В технических, биологических, физиологических и других системах также имеет место обмен информацией, но их учет в общем определении коммуникации слишком сложен.

Вслед за Н. Хомским с 1965 г. коммуникация обычно определяется как общение между *n* лицами, где *n* равно или больше чем два. Коммуникация может состояться и внутри человека (персональная коммуникация), в этом случае она осуществляется с помощью внутренней речи и так называемого кодового языка мышления. Кроме того, коммуникация может состояться и между организациями, государствами и т.п. Тогда говорящий обычно выступает не от себя лично и выражает не свое мнение.

Коммуникация сложна из-за многообразия ее средств. Соответствующая семиотическая шкала начинается с указанного кодового языка мышления, в ней имеются также искусственные, символические и другие языки, естественные языки и языки поэзии, музыки, художественного творчества. В комбинации с устной, письменной или образно-наглядной формой представления этих языков указанная шкала образует многообразный спектр семиотических средств. В ходе речевой деятельности с помощью этих семиотических средств происходит порождение новых и восприятие (интерпретация) имеющихся текстов (объяснений, соответствий и т.п.) о фактах (событиях, явлениях) мира и об идеях, мыслях человека, что и составляет сущность коммуникации. Коммуникация происходит и в случае перевода с одной семиотической системы в другую, при этом факты и идеи остаются теми же. Коммуникация в виде порождения и восприятия текстов-информаций основывается на имеющемся опыте, знаниях, информации и имеет при этом динамический характер, т.е. со временем (в течение коммуникации и при ее повторении) текст-информация может изменяться, дополняться.

На основе приведенного рабочего определения коммуникации можно вывести (см. также [11; 133]) ряд признаков и характеристик коммуникации, которые важны для ПИЯ и тестирования. Коммуникация:

1) является формой социального общения, которая всегда имеет как личностный смысл для одного участника коммуникации, так и общезначимый смысл для обоих участников-партнеров и для общества в целом, причем она протекает в естественной обстановке реальной жизни, поэтому используется аутентичный язык с его возможными

социолингвистическими (диалектными) и стилистическими вариациями;

2) осуществляется между участниками с помощью коммуникативных актов, которые являются ее содержательными, прагматическими единицами, и с помощью речевых действий, которые являются ее операциональными единицами для реализации указанных коммуникативных актов и предшествующих им коммуникативных намерений (коммуникативных задач), при этом ее форма и содержание имеют элементы непредсказуемости и определенной случайности, поэтому используемые речевые действия имеют проблемный, творческий характер, а успешность и результат коммуникации носит вероятностный характер;

3) всегда целенаправлена, ее целью является результативность коммуникации, т.е. обмен информацией (максимально возможная упорядоченность информации), а также сообщение и выявление коммуникативных намерений, т.е. выполнение коммуникативных задач;

4) всегда контекстуальна как в широком экстралингвистическом смысле, т.е. она осуществляется с учетом общественно-политических (идеологических), социально-культурных (исторических) и этно-культурных контекстов, так и в узком, лингвистическом смысле, т.е. она функционирует в связанном тексте, в дискурсе и имеет деятельностный характер, поскольку тексты (в более узком, чем в указанном выше смысле) порождаются в коммуникации и воспринимаются в речевой деятельности;

5) может подвергаться различным «шумам» и ограничениям, поэтому для уменьшения неопределенности и преодоления трудностей в общении партнеры: а) осуществляют вероятностное прогнозирование коммуникативных намерений партнера и оказывают взаимопомощь друг другу при их реализации; б) осуществляют постоянную оценку значимости получаемой и передаваемой информации и используют обратную связь как внутри себя, так и в процессе коммуникации.

Коммуникативная компетенция. Для успешности коммуникации главным условием является наличие достаточной коммуникативной компетенции (КК), а также уменьшение влияния указанных выше ограничений. Для пояснения сущности КК можно исходить из вышеприведенных характеристик коммуникации. Из них следует, что КК состоит из ЛК, которую нужно дополнить следующими компонентами-компетенциями:

1) способностью соединения отдельных предложений в связное устное или письменное сообщение, дискурс, используя для этого различные синтаксические и семантические средства когезии (дискурсивная компетенция);

2) способностью выражать и понимать словосочетания и предложения с такой формой и с таким значением, которые соответствуют (are appropriate) данному социолингвистическому контексту иллокутивного акта коммуникации (социолингвистическая компетенция);

3) способностью эффективно участвовать в общении, выбирая правильную стратегию дискурса, если коммуникации грозит разрыв из-за

шума, недостаточной компетенции и т.п., а также выбирая правильную стратегию для повышения эффективности коммуникации (стратегическая компетенция).

Указанную структуру КК из четырех компонентов предложили в 1980 г. М. Кэнел и М. Свейн [92], частично она была экспериментально доказана Л. Бахманом [88; 89]. Последний предложил еще одну компоненту КК – иллокутивную компетенцию – способность надлежащим образом формировать иллокутивные (речевые) акты (попросить что-то, пригласить, информировать кого-то) в соответствии с ситуацией общения.

Вышеприведенную структуру КК следует дополнить еще одной компетенцией – страноведческой компетенцией, т.е. способностью учитывать в иллокутивных актах общения особенности страны, культуры, истории народа, язык которого изучают и с представителем которого идет общение.

Страноведческая компетенция охватывает и лингвострановедческий компонент, что означает знание и правильное использование таких иноязычных слов и выражений, обозначающих предметы, явления, факты, идеи, которых либо нет в своей стране, либо они называются иначе или менее образно (типа hot dog, foreign secretary).

Рассмотрим, как вышеуказанное отражено в имеющихся в литературе определениях КК. Обычно отмечают, что впервые ввел понятие коммуникативной компетенции Д. Хаймз в 1972 г. [122], связывая лингвистическую компетенцию в понимании Н. Хомского с ее социолингвистической применимостью в разных ситуациях реальной жизни. Р. Валетт [165] считает КК способностью общения в специфических ситуациях. Более точное и развернутое определение дает Н.И. Гез [11]: КК является способностью правильного использования языка в разнообразных социально детерминированных ситуациях. М. Финоккиаро [108] считает КК способностью человека понимать и воспроизводить иностранный язык не только правильно, но и в соответствии с социолингвистической ситуацией реальной жизни, причем речь должна быть соответствующей (appropriate) для определенного говорящего/слушающего/читающего в специфических ситуациях. С. Савиньон [147] приводит в своей книге следующее определение КК: это функциональное владение языком и выражение, интерпретация и обговорение значений (negotiation of meaning), охватывающие общение между двумя и более лицами, принадлежащими к тому или иному речевому сообществу или между одним лицом и письменным или устным текстом.

Определение КК, приведенное в статье И.А. Зимней [17], легло в основу рабочего определения КК в настоящем пособии.

Коммуникативная компетенция – это способность осуществлять речевую деятельность, реализуя коммуникативное речевое поведение на основе фонологических, лексико-грамматических, социолингвистических и страноведческих знаний и навыков и с помощью умений, связанных с дискурсивной, иллокутивной и стратегической компетенцией, в соответствии с различными задачами и ситуациями общения.

3.3. ТЕОРЕТИЧЕСКИЕ ПРОБЛЕМЫ ИЗМЕРЕНИЯ КОММУНИКАТИВНОЙ КОМПЕТЕНЦИИ

Модели КК. Для измерения КК следует иметь ее достаточно четкую модель, которая позволила бы выделить минимально необходимое количество показателей и характеристик КК. После моделей Дж. Манби [134], Г. Уидоусона [169], предложенных в 1978 г., и описанной выше модели КК М. Кэнела и М. Свайна [92] появилась еще модель М. Халлидея [114], признанная в 1982 г. ACTFL (Американским советом преподавателей иностранных языков), модель, предложенная в 1980 г. Р. Клиффордом [100], а также модели, которые предложили в 1983 г. С. Савиньон [147], К. Фаерш и Г. Каспар [106]. Однако все эти модели имеют некоторые недостатки. Модель Дж. Манби слишком сложна, охватывая по анализу потребностей (needs analysis) 256 умений, и подходит скорее для составления подробных программ обучения. В модели Г. Уидоусона недостаточно учтены дискурсивная и стратегическая компетенция, а в двух последних моделях неясно, как и чем разделены компоненты КК и речевые умения.

Рассмотрим вкратце остальные модели.

Модель М. Халлидея исходит из взаимосвязи трех «миров» – текста, его идеи (концепции) и интерпретаторов (interpersonality). Сделано предположение, что «миры» текста и его идеи, т.е. форма и смысл (семантика) предложений текста, имеют другую степень определенности, чем обговариваемый «мир» интерпретаторов, т.е. участников коммуникации. На рис. 5 приведена модель М. Халлидея с переводом ее зон на компоненты КК.

Обозначение зон

по М. Халлидею		по компонентам КК	
1. текстуальная	зона	лингвистическая	компетенция
2. формальная	–"–	страноведческая	–"–
3. идейная	–"–	стратегическая	–"–
4. дискурсивная	–"–	дискурсивная	–"–
5. функциональная	–"–	социолингвистическая	–"–
6. интерперсональная	–"–	иллокутивная	–"–

Области М. Халлидея вне кругов: I – лингвистическая (грамматика, фонология, лексика и т.п.); II – семантическая (понятия, концепции и т.п.); III – прагматическая (деятельность, речевые акты и т.п.).

Модель Р. Клиффорда называют трехсекционной, поскольку владение неродным языком определяется согласно этой модели по трем равнозначным компонентам: функции (способность беседовать, выражать мнение, просить, обсуждать и т.п.), тематике (content/topics) и точности (accuracy) в коммуникативных актах. На рис. 6 эта модель показана треугольником с точками Т (тематика), А (точность) и Ф (функции). Расположе-

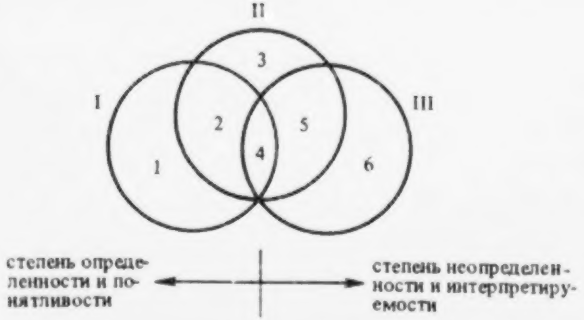

Рис. 5. Модель коммуникативной компетенции М. Халлидея.

ние зон, приведенных на рис. 5, на этом рисунке несколько иное, чтобы ввести верхнюю шкалу коммуникации: факты, идеи – текст, сообщение.

Глобальной КК на рис. 6 называется способность координировать взаимодействие отдельных компонентов КК для обеспечения эффективности и результативности коммуникации, т.е. это вербальные способности человека.

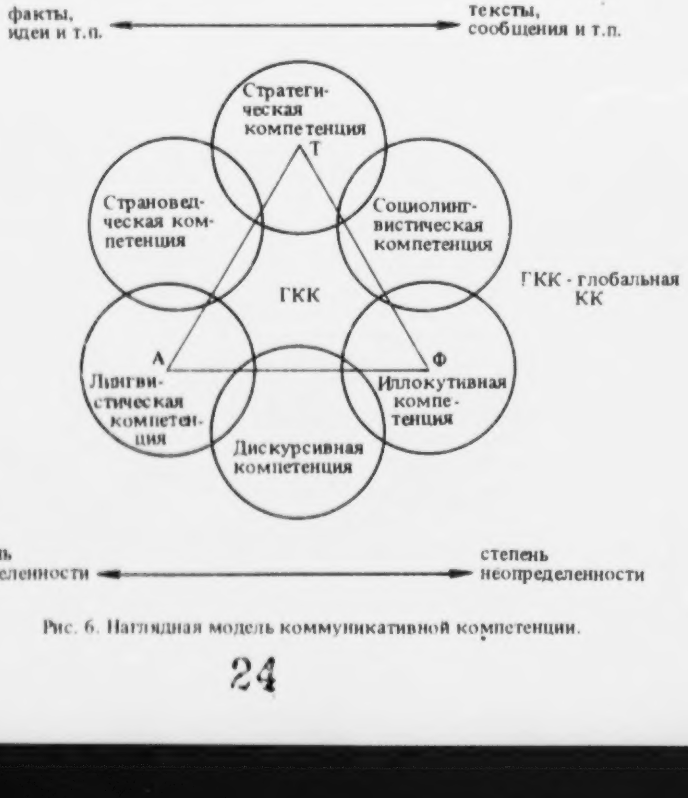

Рис. 6. Наглядная модель коммуникативной компетенции.

Существование отдельной глобальной КК доказано экспериментально в виде генерального фактора в факторном анализе результатов многих языковых и речевых тестов, например, Дж. Оллером [137], а наличие системы из глобального и двух компонентов — лингвистического/прагматического и социолингвистического — Л. Бахманом [88; 89]. Вообще возможны три подхода — только глобальная, унитарная (единая) КК, только компонентная КК (как считалось примерно до 1975 г.) и комбинированная (глобальная КК плюс компоненты). Глобальная КК является, вероятно, общей для родного и неродного языка и во многом, видимо, перекрывается общим интеллектуальным развитием.

Структура факторов, влияющих на измерение КК. Л. Бахман и А. Пальмер [88] объединили все факторы, которые влияют на результат измерения КК, в четыре группы: 1) факторы самих компонентов КК (trait factors); 2) факторы речевых умений (skill factors); 3) факторы метода определения КК (method factors); 4) случайные факторы (random factors), влияющие на КК несистематически.

Первая группа факторов подразделена на организующие компоненты формального использования языка (лингвистическая и дискурсивная компетенция) и на прагматические компоненты (иллокутивная и социолингвистическая компетенция), которые отражают функциональные и социальные характеристики использования языка.

Вторая группа факторов содержит факторы речевых умений, связанные: 1) с видами речевой деятельности; 2) с модальностью представления стимулов и ответов; 3) с формой стимулов и ответов (сознательные-подсознательные, аналитические-синтетические, подготовленные и неподготовленные); 4) со стратегической компетенцией — с общими умениями использовать другие компоненты КК для достижения максимальной эффективности обмена и переработки информации в ходе общения.

Факторы метода определения КК содержат: 1) тип ситуации использования языка — с партнером или без партнера (в языковом тесте); 2) наличие контекста или его отсутствие в заданиях; 3) тип и распределение информации (абстрактная или конкретная, концентрированная или рассеянная соответственно); 4) тип и степень влияния различных ограничений, касающихся тестируемых, тестовой ситуации, организации дискурса, времени и др.

Случайные факторы содержат: 1) психологические факторы (личностные характеристики — мотивация, независимость, устойчивость к неопределенности); 2) факторы взаимодействий факторов из групп 1–3; 3) ошибки измерения.

На основе этой системы факторов Л. Бахман и А. Пальмер определяют КК как состоящую из лингвистической компетенции (ЛК) плюс еще такие компоненты КК, как дискурсивная (ДК), иллокутивная (ИК) и социолингвистическая компетенция (СЛК).

Взаимосвязи между компонентами КК показаны на рис. 7. Теоретически очевидно, что стратегическая компетенция (СК) связана со всеми

остальными компонентами КК, как бы обслуживая их и давая им больше или меньше возможностей для их актуализации. С другой стороны, результат формирования иллокутивных актов подвергается воздействию со стороны всех других компонентов КК.

Рис. 7. Взаимосвязь между компонентами КК.

Указанные взаимосвязи компонентов КК следует проверить экспериментально с помощью корреляционного и факторного анализов.

С точки зрения удельного веса отдельных компонентов КК в общей оценке развитости КК тестируемого следует учесть, что соотношение компонентов КК изменяется в ходе обучения, и наиболее важными, как было показано выше, являются СК и ЛК, за ними следуют ИК и ДК. Несколько менее важными являются социолингвистическая и страноведческая компетенции. Это очень приближенно показано на рис. 8.

Отметим, что в отношении компонентов ЛК имеет место аналогичная картина – наиболее важным является грамматический аспект, но в начале обучения доминируют фонологический и лексический аспекты.

Пять уровней КК на рис. 8 имеют следующее содержание, согласно критериальной шкале ACTFL [100] :

1 – элементарное владение: тестируемый может отвечать на вопросы и задавать вопросы в пределах изучаемого материала и знакомой тематики; он справляется в минимально необходимых ситуациях общения;

2 – ограниченное владение ИЯ для работы: тестируемый может описать ситуацию, рассказать о себе, о своей работе, стране и т.п.; он справляется в знакомых профессиональных и страноведческих ситуациях общения;

3 – минимальное профессиональное владение ИЯ: тестируемый может рассуждать, выдвигать мнения, доказать что-то, говоря о своей работе;

4 – весьма свободное профессиональное владение ИЯ;

5 – тестируемый – практически билингв.

Указанная шкала слишком приближенна, особенно в самом начале, поэтому ACTFL предложил еще промежуточные уровни с плюсами: 0+, 1+, 2+, 3+ и 4+. Из шкалы следует, что студенты неязыковых вузов начинают курс на уровне 1 или 1+ и завершают его на уровне КК 2 или 2+.

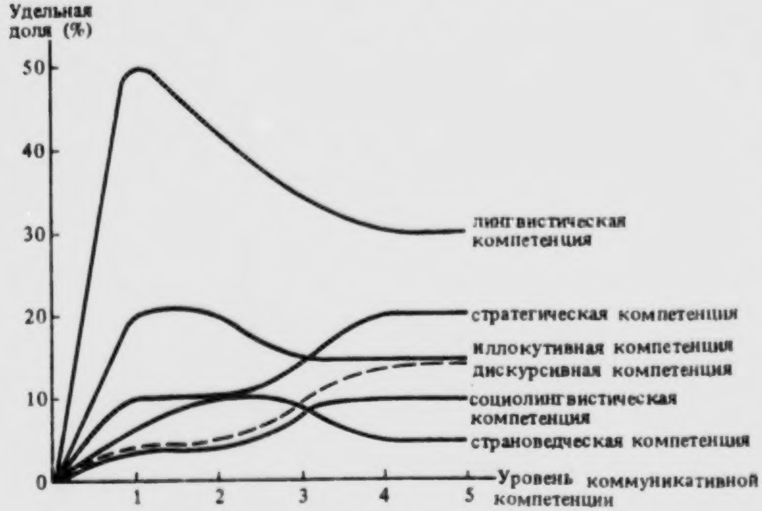

Рис. 8. Оценка динамики соотношения разных компонентов коммуникативной компетенции по мере ее повышения.

В таблице 2 приведена качественная оценка связи некоторых речевых действий с компонентами КК. Из этой таблицы следует, что не всегда обязательно измерять все 6 компонентов КК, как это имеет место при составлении теста диалогической устной речи.

Таблица 2

Качественная связь некоторых речевых действий с компонентами коммуникативной компетенции

Примеры речевых действий	Компетенции, связанные в основном со знаниями			Компетенции, связанные в основном с умениями		
	лингвистическая	социолингвистическая	(лингво)-страноведческая	дискурсивная	иллокутивная	стратегическая
Монологическое говорение	+ +	+	+ +	+	+ +	+
Диалог (говорение и аудирование)	+ +	+ +	+ +	+ +	+ +	+ +

Примеры речевых действий	Компетенции, связанные в основном со знаниями			Компетенции, связанные в основном с умениями		
	лингвистическая	социолингвистическая	(лингво)страноведческая	дискурсивная	иллокутивная	стратегическая
Слушание доклада (аудирование)	+ +	+	+	+	+	–
Конспектирование лекций (аудирование и письмо)	+ +	+	+	+ +	+ +	+
Написание письма, доклада и т.п.	+ +	+	+	+ +	+ +	+ +
Реферирование в ходе чтения (чтение и письмо)	+ +	+	+	+ +	+ +	+
Чтение с определенной целью	+ +	+	+	+	+	–
Чтение доклада (чтение и говорение)	+ +	+	–	+	+	–

Примечание: «+ +» – «очень важно», продуктивное использование компетенций; «+» – «важно», рецептивное использование компетенций; «–» – «менее важно».

3.4. КЛАССИФИКАЦИЯ ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

3.4.1. Языковые и речевые тесты

Как следует из общего определения ЛДТ, их можно подразделить на две группы: языковые и речевые тесты. Языковые тесты измеряют в основном лингвистическую компетенцию тестируемых. Примерами языковых тестов являются лексические и грамматические тесты, как наиболее распространенные в преподавании иностранных языков. Речевые тесты – тесты чтения, аудирования и др. – можно условно разделить в соответствии с применяемой формой контроля на следующие подгруп-

пы: традиционные, прагматические и коммуникативные. Традиционные речевые тесты предназначены для выявления навыков и умений чтения, аудирования и письма с помощью традиционных приемов, например, вопросно-ответных заданий. Примерами прагматических тестов являются тесты восстановления типа клоуз, диктанты и др. Коммуникативные тесты — это в основном тесты устной речи. Все речевые тесты измеряют какие-то компоненты коммуникативной компетенции, сущность которых была рассмотрена выше.

В таблице 3 приведены характеристики четырех степеней коммуникативности ЛДТ. Средние точки (степени) весьма приблизительные и служат скорее для иллюстрации. Как видно из этой таблицы, все виды ЛДТ обладают какой-то степенью коммуникативности, хотя в случае некоторых фонетических тестов она приближается к нулю.

Для практического сравнения степени коммуникативности двух тестов следует суммировать кодированные оценки этих тестов на таких шкалах (по таблице 3), как, например, степень контекстуальности, аутентичности языка, естественности (аутентичности) ситуации тестирования, полноты охвата компонентов или субкомпонентов КК.

Таблица 3

Характеристики степени коммуникативности лингводидактических тестов

№№	Характеристики	Степень коммуникативности ЛДТ			
		минимальная	ниже среднего	выше среднего	максимальная
1.	Активное/пассивное владение ИЯ	пассивное	пассивное	пассивное/активное	активное
2.	Косвенность/непосредственность характера измерения с помощью ЛДТ	косвенный	косвенный	косвенный/непосредственный	непосредственный
3.	Соотношение: лингвистическая форма/семантическое значение	максимально форма	в пользу формы	в пользу значения	максимально значение
4.	Характер психолингвисти-	узнавание, идентификации	соотнесение лингвистичес-	частичное воспроизведение	продуктивное использова-

№№	Характеристики	Степень коммуникативности ЛДТ			
		минимальная	ниже среднего	выше среднего	максимальная
	ческих мыслительных операций при решении заданий ЛДТ	ция лингвистической формы	кой формы со значением		ние языковых явлений
5.	Контекстуальность заданий и ответов	минимальная	относительно высокая	высокая	максимальная
6.	Охват компонентов КК	ЛК	ЛК, страноведческая К	все кроме ИК, СК	все компоненты
7.	Аутентичность тестовой ситуации	нулевая	низкая	средняя	высокая
8.	Аутентичность используемого в ЛДТ языка	низкая	средняя	высокая	максимальная
9.	Степень интерактивности (взаимодействия с партнером)	нулевая	низкая	средняя	максимальная
10.	Доминирующий уровень владения ИЯ, при котором используется соответствующий ЛДТ (шкала ACTFL)	0–1	1–3	2–3	2–4
11.	Представительный пример ЛДТ	фонетический, лексический, грамматический тест	традиционный тест чтения, аудирования	тест типа клоуз, диктант, комбинированный тест	тест-интервью

В таблице 4 приведено сопоставление характеристик усредненных (неэкстремальных) языковых и речевых тестов.

В таблице 4 признаки в п.п. 1–4 относятся к этапу планирования теста, в п.п. 5–8 – к этапу его составления, в п.п. 9–13 – к этапу проведе-

Таблица 4

Сопоставление характеристик языковых и речевых тестов

№№	Признаки и характеристики ЛДТ	Языковые тесты	Речевые тесты
1.	Назначение ЛДТ – определение:	лингвистической компетенции	коммуникативной компетенции
2.	Характер измерения с помощью ЛДТ:	относительно косвенный	относительно непосредственный (в речевой деятельности)
3.	Основная цель применения ЛДТ:	диагноз, успешность обучения	уровень владения речевыми умениями
4.	Подход к составлению:	структурный, аналитический	коммуникативный, комплексный
5.	Характер заданий:	дискретный, вне макроконтекста	интегративный, в макроконтексте
6.	Характер предъявляемого в тесте материала:	несколько искусственный	естественный, в речевой деятельности
7.	Основная концепция шкалы измерения с помощью ЛДТ:	одномерная и узкая	многомерная и комплексная
8.	Структура теста:	обычно с однотипными заданиями	с разнообразными заданиями
9.	Характер ответов на задания ЛДТ:	предопределенный	творческий
10.	Основной тип ответов:	избирательный	со свободно конструируемыми ответами
11.	Процедура проведения теста:	несколько искусственная, групповая	естественная, иногда индивидуальная

№№	Признаки и характеристики ЛДТ	Языковые тесты	Речевые тесты
12.	Состояние тестируемого: 1. роль 2. мотивация	 пассивная заниженная	 активная высокая
13.	Роль тестирующего:	пассивная	активная – партнерская
14.	Подход к оценке результатов теста:	определение ошибок, пробелов	определение того, что сможет и сумеет тестируемый в ходе общения
15.	Процедура оценки результатов теста:	обычно объективная: не зависит от оценивающего	иногда субъективная
16.	Шкала подсчета результатов теста:	относительная: нормы, баллы, проценты	абсолютная и соответствие с определенными критериями
17.	Валидность: 1. внешняя 2. концептуальная (конструктная) 3. сопоставительная	 не очень высокая низкая высокая	 высокая высокая низкая
18.	Надежность:	высокая	не очень высокая
19.	Экономичность:	высокая	низкая

ния теста, в п.п. 14–16 – к проверке результатов теста и в п.п. 17–19 – к интерпретации результатов теста – к показателям качества теста.

Структурный подход к составлению языковых тестов (см. п. 4 табл. 4) используется из-за того, что во время их интенсивного развития и разработки в 50–60-х годах в лингвистике доминировали концепции структурализма и трансформационно-генеративной грамматики. Согласно этим концепциям, основными единицами обучения и тестирования считаются слово, словосочетание и предложение. Тестируемый языковой материал подвергается анализу с целью отбора в тест приемлемых

с разных точек зрения лингвистических единиц. Из-за этого задания языковых тестов имеют дискретный, изолированный друг от друга и от большого (макро)контекста характер. Каждое задание теста измеряет только одну трудность (см. п.п. 7, 8 в табл. 4). Характер предъявляемого в этих тестах языкового материала носит поэтому несколько искусственный характер, что снижает естественность процедуры тестирования, ухудшает мотивацию тестируемых, оставляя им только пассивную, недостаточно творческую роль.

Указанное замечание является причиной того, почему в таблице 4 п. 15 указано, что оценка языковых тестов «обычно» объективна. Дело в том, что объективная, не зависящая от оценивающего оценка результатов теста возможна лишь в случае избирательных тестов, т.е. когда правильный ответ задан среди отвлекающих вариантов ответов (дистракторов). Но составление такого теста по технике множественного выбора (ТМВ) ответов занимает довольно много времени, требует опыта и коллективной работы нескольких составителей. Поэтому при небольших контингентах тестируемых и при однократном применении теста, например, в случае предварительного опробования теста, можно такой тест составить со свободно конструируемыми ответами, т.е. составить тест, имеющий субъективную оценку результатов. Опыт показывает, что ответы на такие тесты можно проверить довольно быстро (соответствующая техника будет рассмотрена ниже).

Оценка результатов языкового теста ориентируется на относительную шкалу, поскольку за основу берут средние результаты данного контингента тестируемых. Результаты речевых тестов могут быть подсчитаны по указанной нормативно-ориентированной шкале, но в случае коммуникативных тестов используется так называемая критериальная шкала, которая была приведена в п. 3.3. По каждому критерию составляют достаточно подробное описание уровней владения ИЯ, обычно отдельно на каждый вид речевой деятельности и на каждый компонент КК.

Показатели качества теста (п.п. 17–19) будут рассмотрены отдельно ниже.

3.4.2. Стандартизованные и нестандартизованные тесты

Стандартизованным тестом (СТ) называется такой тест, который опробован на достаточно большом (500–1000 чел.) и представительном контингенте тестируемых и который имеет стабильные и приемлемые показатели качества, а также спецификацию – паспорт с нормами, условиями и инструкциями для его многократного использования в разных условиях.

Представительность контингента означает, что выборка тестируемых по их данным (возраст, пол, обученность, обучаемость, национальность и др.) представляет такие же данные какой-то более крупной выборки, которую невозможно и нецелесообразно тестировать с помощью

предварительных вариантов СТ. Указанные в определении СТ разные условия использования ЛДТ обычно довольно ограничены и должны быть указаны в спецификации – паспорте.

Нестандартизованными называются преподавательские/учительские (classroom) тесты (ПТ), которые составлены самими преподавателями/учителями для своих учащихся. Они имеют на практике несколько названий – просто тест, тестовая контрольная работа, квиз (quiz). ПТ коротки, и для них необязательно определение всех показателей качества. Однако они отличаются от обычной контрольной работы (задания) тем, что проводится хотя бы одно опробование с целью определения трудности заданий для более обоснованного составления окончательного варианта ПТ.

Имеет смысл ввести еще одну разновидность ЛДТ, промежуточную между СТ и ПТ. Эту разновидность целесообразно назвать локально стандартизованным тестом (ЛСТ). Как вытекает из названия, его отличие от СТ состоит в том, что представительной выборкой является контингент тестируемых одной школы, одного вуза или школ/вузов одного города в пределах 100–300 чел. С одной стороны, это улучшенный, возможно более полный ПТ и, с другой стороны, это менее строгий по показателям качества СТ. Если ПТ составляют 1–2 человека, а ЛСТ – небольшая группа (4–5 человек), то составлением «широко» стандартизованных тестов охвачены целые коллективы.

Таким образом, степень стандартизованности ЛДТ заключается в степени их разработанности, в разной представительности выборок тестируемых для опробования и использования, а также в качестве самих тестов.

В таблице 5 приведена качественная пригодность тестов трех степеней стандартизации для выполнения разных функций тестирования. Стандартизованные тесты будут подробнее рассмотрены в разделе экономичности тестов.

Таблица 5

Выполнение тестами трех степеней стандартизации разных функций тестирования

Функции тестов		Степень стандартизации		
		нестандартизованный (ПТ)	локально стандартизованный (ЛСТ)	широко стандартизованный (СТ)
1. Контролирующая	(успеваемость)	+ +	+	–
2. Диагностическая	(ошибки, пробелы)	+ +	+ +	+

Функции тестов	Степень стандартизации		
	нестандартизованный (НТ)	локально стандартизованный (ЛСТ)	широко стандартизованный (СТ)
3. Успешность обучения (достижения)	+	+ +	+ +
4. Общее владение языком	–	+ +	+ +
5. Прогноз будущей успеваемости обучения	–	+	+ +
6. Обучающая	+ +	+	+
7. Воспитательная	+ +	+	+
8. Исследовательская	+	+ +	+ +

Примечание: «++» – максимально возможное выполнение данной функции; «+» – принципиальная возможность (с недостатками) выполнения данной функции; «–» – невозможность или нецелесообразность выполнения данной функции.

3.4.3. Классификация тестов по целям тестирования

По целям тестирования ЛДТ обычно классифицируются следующим образом (по степени важности):

1. тесты учебных достижений (achievement tests);
2. тесты общего владения иностранным языком (proficiency tests);
3. диагностические тесты (diagnostic tests);
4. тесты способностей (aptitude tests).

Тесты учебных достижений или успешности обучения составляются точно по программе (учебнику) или пройденному языковому материалу и используются для осуществления текущего, промежуточного и итогового контроля. Эти тесты, кроме того, выявляют ошибки и недостатки обучения и поэтому имеют также некоторые диагностические функции. Чем больше охват пройденного языкового материала в тесте, тем лучше проявляются и его диагностические функции. С другой стороны, если составить тест учебных достижений абитуриентов, то, поскольку в него необходимо включить проверку сформированности разных иноязычных навыков и умений, такой тест измеряет и общее владение иностранным языком.

Тесты общего владения ИЯ составляются независимо от программ (учебников) предыдущего обучения. В тест такого рода может быть включена проверка тех иноязычных навыков и умений, которые будут необходимы для дальнейшей работы или учебы. Таким образом, этот

вид теста может выполнять и прогностические функции. Сейчас это все больше тесты для определения коммуникативной компетенции.

Для составления диагностических тестов с помощью предварительного опробования как можно большего количества объектов тестирования выявляются трудности и ошибки, типичные для данного контингента тестируемых. Этот вид теста, естественно, выявляет и качество предыдущего обучения, т.е. является тестом учебных достижений. Э. Инграм [123] считает, что, чем больше контекста в диагностическом тесте, тем хуже выполняется именно диагностическая функция этого теста. Однако это справедливо только в случае языковых тестов.

Прогноз будущей успешности обучения ИЯ может быть составлен на базе теста общего владения ИЯ, если тестируемый уже обучен, но в начале обучения, особенно в самом начале изучения ИЯ используются специальные прогностические тесты, выявляющие разные способности к изучению ИЯ. Такие тесты будут рассмотрены ниже в разделе проблем составления тестов.

Дж. Кларк [98] считает, что диагностические тесты являются разновидностью тестов достижений (diagnostic achievement tests) наряду с тестами общих учебных достижений (general achievement tests). А. Гаррисон [117] выделяет тесты достижений, общего владения ИЯ, диагностические, а также еще так называемые распределительные тесты (placement tests), целью которых является распределение тестируемых по результатам этого теста в примерно равные по уровню обученности учебные группы. Распределительные или вступительные тесты – по существу это либо тесты достижений, если они составлены по программе, либо тесты общего владения иностранным языком.

Р. Валетт [165] также приводит две разновидности тестов достижений, один из которых (тест прогресса) предназначен для текущего и промежуточного контроля успеваемости, а другой (собственно тест достижений) предназначен для итогового контроля.

Из вышесказанного следует, что практически все три основных типа тестов являются многоцелевыми, что показано в таблице 6.

Таблица 6

Характеристики тестов с различными целями тестирования

Характеристика	Название теста		
	учебных достижений (успешности)	общего владения языком	диагностический
Основная цель	определение учебных достижений	определение степени общего владения языком	выявление ошибок и недостатков

Продолжение

Характеристика	Название теста		
	учебных достижений (успешности)	общего владения языком	диагностический
Возможная дополнительная цель	диагностическая	прогностическая	определение достижений развития
Косвенно показывают	степень общего владения языком	учебные достижения	–
Состав	языковые и/или речевые тесты	комплекс из разных видов тестов	в основном языковые
Отношение к программе	точно по программе	независимо от программы	по программе или частично по программе
Наиболее важный вид валидности	по содержанию и сопоставительный	сопоставительный и прогностический	по содержанию
Степень стандартизации	все степени	стандартизованные и локально стандартизованные	преподавательские и локально стандартизованные
Временное отношение	данное состояние и ретроспекция	данное состояние и прогноз на будущее	данное состояние

3.5. ПОКАЗАТЕЛИ КАЧЕСТВА ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

Требования, которые предъявляют к ЛДТ, выражаются в их показателях качества. Многие неудачи, связанные с тестированием в ПИЯ, обусловливаются недостаточным вниманием к показателям качества ЛДТ уже на стадии их планирования. В связи с переходом на коммуникативно-ориентированную методику ПИЯ возрастает роль речевых тестов, а специфические для них показатели качества еще не разработаны.

В классической тестологии (см., например, П. Линерт, А. Анастази и др. [130; 1], обычно указываются следующие основные показатели качества тестов: 1) валидность; 2) надежность; 3) практичность; 4) экономичность.

Валидность. Под валидностью, называемой иногда действенностью, адекватностью, пригодностью, понимается ответ на вопрос: измеряет ли данный тест действительно то, что хотели его составители. В случае языковых тестов ответить на этот вопрос несколько легче, поскольку они однороднее и однозначнее речевых, особенно прагматических и коммуникативных тесты.

В таблице 7 приведены те виды валидности, которые традиционно используются в случае языковых тестов, причем они приведены не по важности, а по степени конкретности их содержания. Первые два вида валидности иногда называют эмпирическими (критериальными), поскольку сопоставительную и прогностическую валидность можно определить количественно через соответствующий коэффициент корреляции. Остальные три вида валидности – по содержанию, конструктную и внешнюю можно определить лишь качественно, через соответствующий анализ тестируемого языкового материала, объектов тестирования, программы и др.

Таблица 7

Характеристика видов валидности ЛДТ

Валидность	Характеристика	Способ определения
Сопоставительная	Насколько хорошо измеряет данный тест то, что хочет тестолог по сравнению с другим тестом или оценкой.	Результаты данного теста сопоставляются (определяется корреляция) с результатами другого теста или оценками преподавателей, которые получены до или сразу после данного теста.
Прогностическая	Насколько хорошо результаты данного теста предсказывают, например, успешность будущей учебы.	Результаты данного теста сопоставляются (определяется корреляция) с результатами другого теста, который проводится через, например, полгода или год учебы.
Содержательная	Представлены ли в данном тесте все важнейшие элементы содержания программы, учебника и т.п.	Проводится логический анализ и сопоставление объектов тестирования с элементами программы, учебника и т.п.
Конструктная (концептуальная)	Насколько объекты тестирования и характер заданий учитывают психолингвистическую модель усвоения рассматриваемого языкового материала	Проводится соответствующий анализ языкового материала теста.

Валидность	Характеристика	Способ определения
	или данную модель коммуникативной компетенции.	
Внешняя	Насколько привлекательными, прагматичными считаются задания теста тестируемыми, преподавателями и администраторами.	Проводится соответствующий анализ языкового материала теста.

При определении сопоставительной валидности важно учесть, что соответствующую корреляцию результатов рассматриваемого теста с результатами другого измерения (теста, оценки) могут снижать:

1) несовпадение объектов измерения данного теста и сопоставляемого теста, а также ранжировки (расположения тестируемых по уровню владения языком) или оценок успеваемости. Например, если сопоставляют результаты лексического теста и оценки успеваемости, то корреляция может быть невысокая из-за того, что в оценках успеваемости учитываются, кроме знания лексики, еще, например, ответы по домашнему чтению или устным темам. Поэтому лучше взять для валидации (определения валидности) рассматриваемого теста какой-нибудь другой тест с известными параметрами (имеющий те же объекты тестирования, достаточную валидность и надежность);

2) временной разрыв между проведением данного теста и сопоставляемого теста. Если этот разрыв более 1–2 недель, то тестируемые могут в какой-то степени «доучиться», поэтому, если данный тест не слишком длинный и есть время на проведение другого, сопоставляемого теста, лучше всего провести оба измерения одновременно.

Для преподавательских тестов необязательно вычисление сопоставительной, а также прогностической валидности.

Конструктная валидность Это степень обоснованности утверждения о том, что данный тест измеряет типичное свойство или состояние, отражаемое в данном конструкте (концепции) [1]. В случае ЛДТ такими концепциями являются психолингвистическая модель усвоения тестируемыми того или иного языкового материала [45], а также модель КК. Более широкими и конкретными теоретическими концепциями являются, например, структурализм, трансформационно-генеративная грамматика, коммуникативная природа речи. Например, тесты, составленные по тестологическим принципам Р. Ладо [127], имеют весьма хорошую конструкторную валидность с точки зрения структурализма.

Роль внешней валидности возрастает с увеличением роли коммуникативных и прагматических тестов. Например, тесты типа клоуз кажутся

тестируемым странными, поэтому следует им объяснить, что в естественном потоке речи, например, при разговоре по телефону на родном языке, мы легко заполняем возникающие «пропуски». Исходя из концепции лингвистической и информационной избыточности языка, нетрудно показать концептуальную валидность тестов типа клоуз, но их внешняя валидность для тестируемых не обязательно будет очевидной.

В случае речевых тестов возникает связанная с внешней валидностью проблема аутентичности, которую нужно понимать не только с точки зрения языка, но и с точки зрения тестовой ситуации, т.е. насколько близкой является данная тестовая ситуация к реальной жизненной. Выше было указано, что достичь полной аутентичности практически почти невозможно.

Надежность. Важным количественным показателем качества ЛДТ является их надежность, под которой понимается определенная стабильность и устойчивость результатов теста. Надежность теста можно выразить через:

1) коэффициент стабильности, который вычисляется как коэффициент корреляции между результатами двух последовательных тестирований одного и того же контингента тестируемых одним и тем же тестом. Такое повторное тестирование не всегда возможно и целесообразно, поэтому чаще прибегают к другим способам определения надежности;

2) коэффициент эквивалентности, который определяется как коэффициент корреляции между результатами двух параллельных (равнотрудных) форм одного и того же теста у одного и того же контингента тестируемых. Однако и в этом случае требуется повторное тестирование, а также составление двух равных форм теста;

3) коэффициент внутреннего постоянства, который определяется как коэффициент корреляции между результатами двух половин одного и того же теста без повторного тестирования. Следует только подсчитать баллы каждого тестируемого (количество правильных или неправильных ответов) отдельно по нечетным и четным номерам заданий теста;

4) коэффициент гомогенности, который определяется с помощью среднеквадратического отклонения результатов теста по формуле Кюдера-Ричардсона. Этот способ наиболее удобен, но в случае малых контингентов испытуемых (менее 50 чел.) с однородным по уровню подготовки составом тестируемых этот способ может дать значительные отклонения от истинной надежности теста.

Основные факторы, влияющие на надежность ЛДТ:

1) репрезентативная выборка объектов тестирования, т.е. максимально возможная покрываемость программы (в случае итогового теста достижений), и репрезентативная выборка контингента тестируемых, т.е. чтобы были представлены как более сильная, так и более слабая части контингента, для которого предназначен данный тест;

2) количество заданий. Чем больше количество заданий ЛДТ, тем надежнее этот тест. Можно пользоваться формулой Спирмена-Брауна для увеличения надежности теста путем увеличения числа заданий:

$$r_t = \frac{m \cdot r_0}{1 + (m - 1) \cdot r_0}, \quad \text{откуда} \quad m = \frac{r_t(1 - r_0)}{r_0(1 - r_t)}, \qquad (1)$$

где r_t – задаваемая, желаемая надежность;

r_0 – исходная надежность;

m – число, показывающее, во сколько раз следует увеличить количество заданий в тесте;

3) трудность заданий. Более трудный тест обычно надежнее легкого при условии, что все задания теста имеют достаточную дифференцирующую способность (ДС). Важнее, однако, иметь в тесте равномерное распределение количества заданий по их трудности, т.е. чтобы задания были с трудностью, например, по проценту правильных ответов в интервале от 30 % до 85 %;

4) степень однородности уровня подготовки тестируемых – при однородном контингенте надежность теста снижается, особенно в том случае, если тест труднее оптимального (50–60 % правильных ответов);

5) дифференцирующая способность заданий теста. Чем она выше в среднем по тесту, тем надежнее тест, причем доля недостаточно дифференцирующих заданий $0{,}30 \leqslant ДС \leqslant 0{,}40$ не должна превышать 10–15 %;

6) стабильность внешних условий тестирования – устных инструкций, мотивации тестируемых, времени, отводимого на тестирование и др.;

7) стабильность и стандартизация оценки субъективных речевых тестов.

Практичность теста. Под практичностью ЛДТ понимают следующие его качества:

1) доступность инструкций теста и содержания заданий теста для понимания тестируемыми; 2) сравнительную простоту организации тестирования; 3) возможность проведения теста в обычных для школы или вуза условиях; 4) сравнительную простоту проверки опросных листов, подсчета результатов теста и подведения итогов.

Проблема экономичности тестов рассматривается ниже отдельно.

Дополнительные показатели качества лингводидактических тестов. Длительная практика автора по составлению разных стандартизованных тестов показала желательность использования следующих дополнительных показателей качества этих тестов:

1) репрезентативность; 2) степень совпадения средней трудности данного теста с оптимальной средней трудностью; 3) вариация средней трудности; 4) средняя дифференцирующая способность теста; 5) доля недостаточно дифференцирующих заданий.

Репрезентативность контингента ЛДТ показывает, насколько представительным был контингент для опробования предтеста по сравне-

нию с основной выборкой или генеральной выборкой. Если тест не обладает репрезентативностью контингента, то при его использовании может ухудшаться его качество (падает валидность и надежность).

Репрезентативностью объектов тестирования называется покрываемость программного материала (или его части) его выборкой в тесте. В тест следует включать наиболее важные и представительные элементы и части пройденной программы.

Чем больше степень стандартизации теста, тем больше необходимость и возможность для составителей теста регулировать его среднюю трудность. Регулировать нужно, однако, всегда в сторону так называемой оптимальной средней трудности теста, чтобы, с одной стороны, дать и более сильным тестируемым достаточно возможностей проявить себя, а, с другой стороны, избежать такого положения, что более слабый по уровню подготовки тестируемый сможет правильно выполнить лишь небольшую долю заданий теста.

Одной из целей предварительного опробования тестов как раз и является оптимизация его средней трудности.

Если принять допустимый интервал трудности задания, например, по проценту правильных ответов в пределах от 35 % до 85 %, то оптимальной трудностью следует считать 60 % для широко стандартизованных тестов. Это согласуется с обычной средней вариацией результатов этих тестов в пределах 15 %, и тогда в случае нормального распределения результатов тестов 95 % тестируемых выполнят тест с результатами в пределах от 30 % до 90 %. Это значительно уменьшает возможность большого числа стопроцентных результатов, а также слишком низких результатов в пределах границ случайного угадывания (25 %).

Оптимальная средняя трудность преподавательских тестов значительно выше указанного, поскольку эти тесты выявляют успеваемость учащихся. Например, официальная шкала Минпроса ЭССР предлагала ставить оценку «5» за выполнение теста или контрольной работы в объеме 95–100 %, оценку «4» – за 75–94 %, «3» – за 50–74 % и «2» – за 10–49 %.

Вариация выражается отношением среднеквадратического отклонения результатов теста к среднему арифметическому. Она полезна для оценки надежности, для характеристики гетерогенности разных контингентов тестируемых и для способности данного теста «растягивать» данные об уровне подготовки тестируемых.

Дифференцирующая способность (ДС) задания является, наряду с его трудностью, важной характеристикой задания. При «очистке» предварительного теста от некачественных заданий обычно можно ориентироваться на трудность заданий.

За пределами указанного интервала трудностей от 30–35 % до 80–85 % рассматриваемое задание уже не может в достаточной мере дифференцировать сильных и слабых тестируемых.

Часто, однако, качественных заданий (ДС больше 0,40) не хватает и в тест приходится включать задания с ДС в интервале от 0,30 до

0,40. Доля таких заданий у отечественных и зарубежных тестов колеблется в пределах от 10 % до 20 % [24].

В таблице 8 приведены типичные количественные значения основных и дополнительных показателей качества ЛДТ.

Таблица 8

Количественные значения показателей качества лингводидактических тестов

Показатели качества	Разновидность ЛДТ		
	преподавательский	локально стандартизованный	стандартизованный
1. Сопоставительная валидность	–	0,80–0,85	0,85–0,90
2. Прогностическая валидность	–	0,60–0,70	0,70–0,80
3. Надежность	0,60–0,80	0,75–0,85	0,85–0,95
4. Оптимальная средняя трудность теста	70–80 %	60–70 %	55–65 %
5. Вариация	8–12 %	10–15 %	10–20 %
6. Средняя дифференцирующая способность (одной части теста)	0,40–0,45	0,50–0,55	0,55–0,60
7. Максимально допустимая доля недостаточно дифференцирующих заданий	20–30	15–25	10–20

3.6. ЭКОНОМИЧНОСТЬ ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

Экономичность ЛДТ имеет большое значение в первую очередь для стандартизованных тестов, поскольку эти тесты рассчитаны на большое количество тестируемых и на их составление затрачивается много сил и времени. Вопросы экономичности ЛДТ следует особенно учитывать на этапах их планирования и составления.

В литературе по тестам одни авторы вообще обходят вопросы об экономичности, хотя эти вопросы желательно учитывать при составле-

нии и преподавательских тестов. Другие авторы ограничиваются лишь короткими определениями экономичности, считая вытекающие из проблем экономичности вопросы легко разрешимыми. Например, Р. Ладо [127] считает экономичным тест, в случае которого можно положительно ответить на вопрос: измеряет ли данный тест то, что хотят его составители, за разумное время, учитывая условия тестирования? Интересно то, что Б. Керрол [95], у которого совсем другой подход к ЛДТ, в этом отношении задает практически аналогичный вопрос: дает ли тест так много информации, как необходимо, с минимальными расходами времени, усилий и ресурсов? Некоторые аспекты экономичности ЛДТ освещены в статьях автора настоящего пособия [9, 34, 24].

Экономичным следует считать такой ЛДТ, который при заданных условиях – цели, объекты тестирования, возраст и уровень подготовки тестируемых – дает максимум достоверной информации о тестируемых с минимальными затратами времени и усилий на его составление, проведение и обработку. Из этого определения следует, что составление экономичного теста требует оптимизации ряда показателей теста при решении следующих проблем:

1) увеличение числа заданий теста повышает его надежность, но в то же время понижает его экономичность, поскольку требуется больше сил и времени на их составление и выполнение;

2) чем выше степень коммуникативности ЛДТ (измерение более контекстуально и менее косвенно), тем выше его валидность, но объем теста также больше, и тест менее экономичен.

Ниже рассмотрены такие показатели экономичности ЛДТ, как приблизительное время, затрачиваемое на составление его заданий и на проведение разных ЛДТ, количество заданий (объем) ЛДТ, среднее время, затрачиваемое на выполнение однотипных заданий конкретного типа, средняя длина заданий в словах и приблизительное время, уходящее на проверку заданий ЛДТ. Такие данные способствуют конкретизации этапов разработки тестов и особенно этапа планирования как составления, так и объема самого ЛДТ.

Приблизительное время, затрачиваемое на составление заданий ЛДТ, зависит от опыта составителей, типа заданий, методики работы и наличия подходящих вспомогательных материалов – текстов, иноязычных энциклопедий, словарей, пособий. В случае поиска в оригинальной литературе текста-стимула микроконтекстных заданий, как, например, было в случае тестов серии АТУ, на одно задание требовалось до двух часов. В случае поиска текста заданий в отечественных учебниках и пособиях на одно задание (по опыту работы с тестами серии ИТА) потребовалось 0,75–1,5 часа, а в случае написания текста задания самими составителями (в тестах по грамматике часто поступают именно так) – 0,5–1 час на одно задание. Время, необходимое для составления 3–4 дистракторов в случае выборочных ответов, также зависит от опыта составителей и методики работы (индивидуальная, групповая или смешанная). В среднем на составление дистракторов может потребоваться от 20 минут

до одного часа на одно задание, которое следует прибавить ко времени на составление (поиск) текста заданий.

Время на проведение (выполнение) стандартизованных тестов обычно ограничено (см. табл. 9). В случае тестов серии ИТА вариация во времени предтеста ИТА-82 была весьма большая, а корреляция между временем выполнения и результатом теста не превышала величины 0,4–0,5 у отдельных вузов, поэтому в окончательном варианте ИТА-84 время выполнения теста не фиксировано. Ограничение времени, однако, желательно сделать вузам – пользователям этого теста – в зависимости от своих условий.

Согласно данным таблицы 9 время на проведение стандартизованных тестов колеблется от 240 до 75 минут (ИТА-84) в зависимости от объема (количества заданий), а также от трудности заданий теста, характера объектов тестирования и типа заданий. По сравнению с указанными факторами на время выполнения теста несколько меньше влияют качество заданий и отвлекающих ответов, установка и отношение тестируемых к тесту и его выполнению, общая психологическая атмосфера при проведении теста (важность результатов теста, ограниченность во времени, приближение перерыва и др.). Время, отводимое на тест, можно сэкономить путем оптимизации его объема – доведения количества заданий по формуле Спирмена-Брауна (см. раздел о надежности тестов) до заданного уровня. Для этого нужно провести опробование предтеста.

Таблица 9

Показатели качества и экономичности некоторых стандартизованных тестов

Тест и его части	Количество тестируемых	Показатели качества			Экономичность	
		средняя трудность %	сопоставительная валидность	надежность	количество заданий	время, мин.
ТОЭФЛ – США /43/ (1981)	100 000 каждый год	50,1	0,87–0,89	0,94	150	120
Аудирование		51,5	–	0,88	50	40
Грамматика и письмо		48,6	–	0,84	40	25
Чтение и лексика		49,4	–	0,89	60	55
СЛЕП – США /44/ (1980)	6000	74	0,57	0,96	150	85
1. Аудирование		90	–	0,89	25	–
2. Диктант		89	–	0,74	20	–
3. Пояснение карты		85	–	0,81	12	–
4. Слушание беседы		78	–	0,71	28	I–IV 40
5. Чтение (картинки)		74	–	0,72	12	–

Продолжение

Тест и его части	Количество тестируемых	Показатели качества			Экономичность	
		средняя трудность %	сопоставительная валидность	надежность	количество заданий	время, мин.
6. Сопоставление картинки)		68	–	0,84	15	–
7. Клоуз-тест по ТМВ		62	–	0,89	40	V–VIII
8. Чтение		48	–	0,69	8	45
ДЕЛТА – ФРГ /45/ (1977)			0,51–0,60		240	240
1–3. Лексика	153	58	–	0,79	55	–
4–6. Грамматика	153	57–72	–	0,45– –0,80	55	–
7. Диктант	227	70	–	–	50	–
8. Синтаксис	127	45	–	0,70	20	–
9. Клоуз-тест	127	79	–	0,68	30	–
10. Перевод предложений	127	56	–	0,67	30	–
ИЭА ЮНЕСКО (1972)	8049				169	160
Аудирование /129/		46,1	0,81	0,75	36	30
Чтение		43,0	0,82	0,77	60	60
Лексика		25,7	–	0,56	16	–
Грамматика		50,0	–	0,81	8	–
Чтение		51,7	–	0,82	36	–
Письмо (сочинение)			0,82–0,86	0,75	31	40
Говорение	311		0,73	0,79	42	30
ИТА-83, I ур. [115] СССР, тест вузовского исходного уровня, английский язык	3249	64,4	0,81–0,89	0,87– –0,90	135	82
Лексика		62,2	0,81	0,89	50	–
Грамматика		62,8	0,61	0,84	25	
Чтение		56,0	0,50	0,74	10	
Клоуз-тест		65,0	0,79	0,84	50	
ИТА-83, II ур. [115] на материале VI–VIII классов	4127	54,3	0,74–0,86	0,85– –0,89	135	91
Лексика		62,4	0,70	0,87	50	
Грамматика		55,7	0,64	0,83	25	
Чтение		67,0	0,40	0,70	10	
Клоуз-тест		55,0	0,76	0,81	50	
АТУ-86 [116], СССР, тест успешности II курса						
Лексика	2107	50,3	–	0,88	72	–
Клоуз-тест	1590	51,3	–	0,86	50	–

Тест и его части	Количество тестируемых	Показатели качества			Экономичность	
		средняя трудность %	сопоставительная валидность	надежность	количество заданий	время, мин.
Грамматика	1120	48,4	–	0,89	27	–
Стандартизованный тест по грамматике для II курса /105/ (1976)	примерно 560	54,0	0,68	0,75	25	30
Национальный тест английского языка для шведских гимназий (1985)	3400	61,2	–	–	100	105
Лексика		54,0	–	0,78	18	35
Клоуз-тест со СКО		61,4	–	0,90	35	
Чтение (длинный текст)		71,7	–	0,67	12	35
Чтение (короткий текст)		56,7	–	0,73	10	
Лексико-грамматический тест		58,6	–	0,74	14	35
Аудирование		70,6	–	0,53	11	

Зависимость времени выполнения теста от количества заданий выражается через среднее время, затрачиваемое на выполнение заданий одного типа. Последнее рассмотрено несколько ниже. Зависимость времени выполнения теста от трудности заданий можно выразить через соответствующий коэффициент корреляции, который при количестве заданий порядка 30–40 и более или менее оптимальной трудности теста имеет порядок 0,20, т.е. статистически незначимую величину (см. табл. 10). Однако при увеличении объема теста до более 60–80 заданий с этой зависимостью следует считаться, так как более сильные тестируемые заканчивают тест быстрее, и можно время на выполнение теста ограничить. Как видно из таблицы 10, у тестов типа клоуз имеется более отчетливая связь между временем и результатами теста.

Количество заданий теста определяется обычно соображениями надежности теста, а общий объем комплексного теста зависит еще от необходимости включения в тест множества объектов тестирования.

Таблица 10

Экспериментальные данные об экономичности разных типов заданий

Тип заданий ЛДТ	Среднее время		Коэффициент корреляции: трудность теста – время на его выполнение
	на выполнение 1 задания, мин./зад.	на проверку 1 задания, сек./отв.	
I. Внеконтекстные лексические задания			
1. Перевод слов по СКО (свободно конструируемый ответ)	0,2–0,4	3–6	0,28–0,30
2. Сопоставление списков слов родного языка – ИЯ по ТМВ (технике множественного выбора)	0,3–0,6	1–2	0,19–0,23
3. Группировки по ТМВ	0,4–0,5	1–2	0,20–0,22
II. Микроконтекстные задания			
4. Завершение по ТМВ	0,5–0,8	1–2	0,18–0,20
5. Синонимы или парафраза по ТМВ	0,6–0,9	1–2	0,11–0,18
III. Макроконтекстные задания			
6. Чтение, ответы на вопросы по ТМВ	0,8–1,2	1–2	0,11–0,15
7. Тест типа клоуз по ТВМ	0,4–0,8	1–2	0,3–0,4
8. Тест типа клоуз со СКО	0,3–0,7	3–6	0,4–0,6
9. С-тест и Б-тест	0,2–0,4	3–6	–
10. Гибкий Б-тест	0,3–0,5	3–6	–
11. Аудирование, ответы на вопросы по ТМВ	0,6–0,9	1–2	–
12. Аудиоклоуз по ТМВ	0,2–0,4	1–2	–

48

Тип заданий ЛДТ	Среднее время		Коэффициент корреляции: трудность теста — время на его выполнение
	на выполнение 1 задания, мин./зад.	на проверку 1 задания, сек./отв.	
13. Аудиоклоуз со СКО	0,15–0,3	3–6	–
14. Диктант (100 слов), чтение трижды	4–6 мин.	10–20 слов/мин.	–
15. Письменный перевод текста со словарем	20–30 печ. зн./мин.	100–200 печ. зн./мин.	0,2–0,5

Из данных таблицы 9 следует, что максимальный объем стандартизованных тестов составляет 130–160 заданий. В составе комплексных тестов максимальные объемы частей лексики составляют 40–60 заданий, по грамматике – 25–40 заданий, по чтению – 20–30 заданий и столько же – по аудированию. По данным Р. Тейлора и Д. Оллера младшего объем клоуз-тестов первоначально составлял не менее 50 заданий, но Л. Бахман [87] предполагает, что при выборочном методе составления заданий-пропусков в тексте можно ограничиться 30–40 заданиями без ущерба надежности такого клоуз-теста.

Среднее время на выполнение однотипных заданий ЛДТ зависит от типа, длины и трудности заданий, от характера отвлекающих и от других факторов. Время на проведение теста, указанное в инструкциях к тестам, обычно заметно выше реального среднего времени на выполнение теста. В случае тестов на чтение и аудирование время на чтение и слушание текста входит в приведенные значения темпа работы. В этих случаях приближенными расчетными значениями можно принять 1 мин./зад. для чтения, 0,8 мин./зад. для аудирования и 0,6–0,7 мин./зад. для лексических и грамматических заданий с выборочными ответами (ТМВ).

Средняя длина заданий в словах без учета пропуска в заданиях типа завершения приведена в таблице 11, откуда видно, что это среднее значение мало зависит от объектов тестирования. Дело в том, что вариация в пределах одного теста весьма велика – бывают задания длиной в 3–4 слова и 15–20 слов на одно предложение. Вариация средней длины от теста к тесту несколько меньше: от 6,6 слов в тесте ИЭА [129] до 14,0 в Мичиганском тесте [163] в случае тестов по грамматике и от 6,8 слов (ИЭА) до 12,3 в тесте ТОЭФЛ [161] в случае тестов по лексике. Длина заданий зависит от уровня подготовки контингента тестируемых, кото-

Таблица 11

Средние данные об экономичности зарубежных и отечественных тестов

Показатели	Аудирование	Клоуз-тест	Чтение		Лексика		Грамматика	
			отечественная	зарубежная	отечественная	зарубежная	отечественная	зарубежная
Общее количество проанализированных заданий	165	1247	138	62	746	330	1014	176
Общее количество проанализированных текстов	55	29	30	16	–	–	–	–
Среднее количество текстов в одной форме	4	1	3	4	–	–	–	–
Количество заданий на 1 текст	3	43	4–5	5	–	–	–	–
Средний объем текста (печ. зн.)	645	2000	670	640	–	–	–	–
Средний объем текста (печ. зн.) на 1 задание	215	47	145	128	–	–	–	–
Средняя длина задания в словах	9,5	9,0	9,7	10,2	8,0	9,4	9,0	10,0
Среднее время на 1 задание	0,75 11 сек. на ответ	0,8	–	0,9	0,65	0,8	0,70	0,8

рым предназначен тот или иной тест. Средняя длина заданий имеет довольно высокую корреляцию со временем на их выполнение: коэффициент 0,69 для 23 тестов (средняя длина – среднее время). Корреляция средней длины заданий с трудностью заданий имела порядок 0,25–0,39 в случае тестов серии ИТА. Любопытно, что, чем длиннее задание по грамматике, тем оно легче, а в случае лексических заданий наоборот. Это

сигнал о том, что средняя длина 6,0–6,8 слов в тесте ИТА-84 (I и II уровни) слишком мала. Похоже, что приведенное в таблице 11 среднее значение 9 ± 1 слово и является оптимальным с точки зрения создания в тестовых заданиях достаточно адекватной лингвистической и экстралингвистической ситуации.

Время на проверку заданий со СКО зафиксировано в условиях, когда варианты ответов тестируемого сопоставляют с правильными ответами на трафарете, который помещен как можно ближе, лучше слева, к колонке порядковых номеров на листе ответов. Время на проверку заданий по ТМВ зафиксировано при проверке с помощью трафарета с окнами на местах правильных ответов. Обычно трафарет окрашен, и места правильных ответов соединены жирной линией для направления движения глаз при проверке.

4. КРАТКАЯ ИСТОРИЯ РАЗВИТИЯ ЛИНГВОДИДАКТИЧЕСКОГО ТЕСТИРОВАНИЯ

4.1. РАЗВИТИЕ ТЕСТИРОВАНИЯ ЗА РУБЕЖОМ

Введение. Ознакомление с историей развития ЛДТ позволяет лучше понять его нынешнее состояние и перспективы на будущее. Оно показывает, что существуют достаточно четкие периоды и направления, которые связаны с развитием теории и практики педагогики, психологии, лингвистики и методики преподавания ИЯ, а также с особенностями различных стран и регионов в области образования и социально-экономического строя.

Настоящий обзор основывается на материалах брошюры И.А. Цатуровой [77], монографии И.А. Рапопорта и др. [62], монографии Л. Бубениковой и др. [91], а также статей А. Дейвиса [103, 104], В. Уайтсона [168] и Дж. Кларка [99][1]. В большинстве из указанных выше монографий не приведено, кроме того, достаточно полного теоретического обоснования используемого в них подхода к составлению ЛДТ, т.е. они имеют скорее характер справочных пособий по практике и технике тестирования.

Периодизация. Первые языковые тесты (в современном понимании) составлялись в 1925 г. в США и Канаде. Однако ЛДТ до сих пор теоретически и практически тесно связаны с педагогическими и психологическими тестами, поэтому целесообразно коротко рассмотреть историю

---

[1] В основных монографиях по ЛДТ, например, таких авторов, как Р. Ладо [127], Р. Валетте [165; 166], Д. Харрис [116], А. Дейвис [104], Дж. Кларк [98], Дж. Хитон [118; 119], Д. Аллен [84], Б. Керрол [93; 95], А. Гаррисон [117], Х. Мадсен [132], М. Финоккиаро, С. Сако [108] и в других не приведен обзор развития ЛДТ.

дидактического контроля и психологических тестов до 1925 г. Для этого можно ввести следующие периоды развития тестирования:

1) Дотестовый период дидактического контроля. Этот период начинается примерно в 2200 г. до н.э. в Китае введением регулярных проверок-экзаменов знаний и умений чиновников. Ф. Дюбуа [105] называет эти экзамены тестами, что однако вряд ли можно считать правильным. В 1115 г. до н.э. в Китае была введена система экзаменов, проверяющих навыки в письме, арифметике, музыке и других предметах для кандидатов, желающих получить административную должность в госаппарате.

Система экзаменов, введенная в Китае в 1370 г., была трехступенчатой: на первой ступени проверялась каллиграфия, на второй (уровень провинции) – экзамен длился 3 дня и одну часть экзамена – сочинение – переписывали клерки для исключения влияния каллиграфии и субъективности столичных экзаменаторов. Указанную систему отменили только в 1905 г. после введения новой системы образования. Ее положительными чертами являлись относительная объективность оценки, открытые и равные условия для всех кандидатов, а также устойчивость, проверенная временем. Недостатками являлись узкая социальная база (не все могли десятками лет заниматься подготовкой к экзамену), большая неграмотность населения, отсутствие контактов с другими культурами.

До середины XIX века дидактический контроль в ПИЯ осуществлялся в зависимости от применяемой методики ПИЯ, в основном путем письменных переводов или сочинений, иногда и путем устного опроса.

2) Период зарождения педагогических и психологических тестов (с 1864 г. по примерно 1890 г.). В 1864 г. англичанин Дж. Фишер составил и опробовал первые педагогические тесты [77], но его идеи остались на десятилетия неиспользованными. Только бурное развитие капитализма во второй половине XIX века и связанное с этим ускорение развития науки и образования потянули за собой развитие психологии, лингвистики и начался период реформы в методике ПИЯ. Однако ни натуральный и прямой, ни комбинированные методы того периода не использовали по понятным причинам тесты. Нужна была соответствующая психометрическая теория.

3) Период преимущественного развития психологических тестов. В 1890 г. Дж. Кеттел [96], американский психолог, ввел понятие «тест» как способ психологических экспериментов. В 1894 г. Дж. Райс составил педагогические тесты для определения степени сформированности различных навыков [77].

Большое значение для развития педагогических и психологических тестов имели исследования Р. Торндайка и изданная в 1904 г. его книга «Введение в теорию психологических и социальных измерений». Под его руководством в 1908 г. был составлен первый стандартизованный педагогический тест [1]. С его именем связывают также направление измерения уровня умственного развития [162]. По его периодизации с 1915 г. по 1930 г. произошел бум развития психологических тестов и завершилась работа над созданием психометрической теории [1].

Из-за альтернативных возможностей, консерватизма и большей социальной ответственности педагогические тесты по своей разработанности стали отставать от психологических тестов. Поэтому второй период условно назван периодом доминирования психологических тестов.

4) Начальный период ЛДТ (с 1925 г. по примерно 1955 г.). В своей периодизации ЛДТ Б. Спольски [153] назвал этот период донаучным по той причине, что недостаточное внимание уделялось качеству тестов, особенно их валидности и надежности. Первый тест по ИЯ составил в 1925 г. Б. Вуд. Его тесты использовались на выпускных экзаменах нью-йоркских школ.

Эти тесты имели в своем составе части лексики, грамматики и чтения на французском и испанском языках. Важной вехой в развитии ЛДТ была книга В. Хенмона [120], посвященная исследованиям с 1924 по 1927 г. Американским и Канадским комитетами по изучению современных языков (modern languages). В этой книге изложены все важнейшие принципы составления, проведения и обработки результатов ЛДТ, которые действительны и в настоящее время. Примерно в это же время появляются первые работы по разработке прогностических тестов (тестов лингвистических способностей).

5) Структурно-психометрический период (с 1955 г. по примерно 1970 г.). Такое название этому периоду было дано Б. Спольским [153] из-за того, что для составления тестов использовался структурный подход в лингвистике и достижения психометрической теории.

ЛДТ предыдущего периода недостаточно учитывало достижения лингвистики и методики ПИЯ, тесты разрабатывались по примеру психологических тестов. Этот недостаток постепенно преодолевался после второй мировой войны с появлением аудиолингвального и позже аудиовизуального методов. Структурно-психометрический период начинается в 1956 г. с выходом в свет книг Ф. и Э. Шоннелов [148] и Дж.Р. Гербериха [111], в которых разрабатываются все важнейшие теоретические и практические вопросы ЛДТ. В начале этого периода был составлен тест способностей Дж. Керрола (MLAT) [94]. Однако, фундаментальным трудом следует считать монографию Р. Ладо [127]. Р. Ладо исходил из тесного взаимодействия обучения и тестирования, предлагая включать в тест те элементы языка, которые по контрастивному анализу являются трудными для обучающихся. Он считал, что знание этих трудностей, умение их преодолеть дает возможность изучить язык. Дж. Керрол уже в 1961 г. назвал этот подход дискретным, так как им в основном предполагалось тестировать в одном предложении-задании одну проблему (трудность). Хотя в отношении чтения и аудирования подход Р. Ладо не был сугубо дискретным, многие способы и приемы тестирования Р. Ладо применимы и сейчас в языковых тестах. Положительным в его работах было то, что он поднял на должную высоту разработанность и научность теста и техники тестирования, а также требования достаточной надежности и валидности ЛДТ.

Влияние положений, изложенных в монографии Р. Ладо, испытывают практически все авторы монографий по ЛДТ вплоть до 1980 г.

В этот период были составлены известные и используемые до сих пор стандартизованные тесты типа TOEFL (Test of English as a Foreign Language), Мичиганский тест и др. в США, EPTB (English Proficiency Test Battery) и ELBA (English Language Battery) в Англии и др. [165]. В США была создана специальная Служба педагогического тестирования (ETS), работающая на коммерческой основе, для разработки и проведения стандартизованных тестов типа ТОЭФЛ, которые проделывают каждый год сотни тысяч иностранных кандидатов для поступления в вузы США.

6) Современный период, который Б. Спольски назвал психолингвистически-социолингвистическим из-за того, что в конце 60-х годов социолингвисты обратили внимание методистов на то, что обучение и тесты на иностранном языке не соответствуют реальным жизненным ситуациям с их возможными социолингвистическими вариациями. Обобщая, предыдущий период можно назвать аналитическим, а современный – синтетическим. В лингвистике, например, М. Хомски и др. сделали ошибку, отделяя для упрощения грамматику от семантики, а затем, например, Катц, Фодор и др. совершили аналогичную ошибку в семантике, считая язык самостоятельной системой вне речевой деятельности, без прагматического аспекта. В 70-е годы методисты стремились использовать синтетический, прагматико-коммуникативный подход.

А. Дейвис приводит нижеследующее сравнение признаков двух указанных периодов [103]:

Структурно-психометрический период	*Психолингвистически-социолингвистический период*
1. аналитический подход	синтетический подход
2. дискретные задания	интегральные задания
3. ориентированная на нормы оценка	ориентированная на критерий оценка
4. рецептивные виды речевой деятельности	продуктивные виды речевой деятельности
5. лингвистическая форма	коммуникативная функция
6. лингвистическая компетенция	коммуникативная компетенция
7. поверхностная структура (микроконтекст)	глубинная структура (макроконтекст)
8. главное – надежность	главное – валидность
Добавим: 9. – объективность	– аутентичность

В рамках этого подхода можно выделить следующие направления тестирования, приведенные в исторической последовательности:

а) Использование интегративных тестов на основе теории (грамматики) вероятностного прогнозирования (expectancy grammar), которую

предложил Дж. Оллер-младший [136; 137] в начале 70-х годов. Это тесты типа клоуз, диктант и др. В этих тестах используется то обстоятельство, что обычная избыточность информации в тексте, сообщении и т.п., достигающая около 50% в английском языке, уменьшена за счет пропусков в тексте (тест типа клоуз) или за счет прерываний чтения текста в диктанте.

б) Направления тестирования, пытающиеся соединить аналитический, структурный подход с новым коммуникативно-ориентированным подходом. Например, в экспериментальных исследованиях коммуникативной компетенции, проведенных в 1972 г. С. Савиньоном [145], А. Пальмером [140] и Х. Коморовской [126] в 1979 г. и др., были использованы те же показатели для КК, которые были предложены в предыдущем периоде для тестов устной речи, т.е. лексика, грамматика, беглость речи, произношение, понимание и др. К ним прибавили такие новые показатели, как количество и качество коммуникации, понятливость, уместность (appropriateness) и др. В 1978 г. Дж. Кларк [97] издал сборник статей под названием «Прямое тестирование навыков говорения», посвященных этим направлениям.

в) Функционально-понятийное направление, которое было предложено для составления программ обучения Уилкинсом [170] в 1976 г., а также рабочей группой Европейского Совета (см., например, Ван Эк и др. [167]). Это направление использовал в тестировании Б. Керрол в 1980 г. [95]. М. Финоккиаро [107] считает, что коммуникативные функции можно разбить на следующие группы: персональные, межличностные, направляющие (иллокутивные акты убеждения, просьбы), референциальные (ссылки на прошлое, настоящее и будущее действия), воображения (сочинять рассказ, стихотворение и др.). В соответствии с коммуникативной ситуацией иллокутивные акты формируются с помощью тех или иных понятий (notions), т.е. лексики в соответствующей форме. Для составления соответствующих тестов следует сначала провести анализ основных, обобщенных потребностей тестируемых (выбор среди изученных коммуникативных функций) и разработать соответствующую критериальную шкалу.

г) Гибридное направление тестирования, предложенное А. Омаггио в 1983 г. [139] и описанное вкратце Дж. Кларком в его обзоре [99]. Как указывает название, это направление также стремится сочетать положительное из структурного направления с новыми взглядами на язык: «гибридные тесты качественно сочетают грамматику и контекст, структуры и ситуацию». Причиной использования гибридного подхода следует считать недостаточную разработанность как теории коммуникативно-ориентированной методики обучения, так и теоретически и экспериментально доказанной модели КК. Это направление приподнимает на должную высоту среди других компонентов тестирование лингвистической точности.

Большое влияние оказала на тестирование устной речи в США соответствующая методика интервью, разработанная уже в середине 50-х

годов Институтом МИД США (FSI). С помощью стандартизации процедур проведения интервью и, главное, его оценки стали приемлемыми коэффициенты надежности (стабильности) рейтеров, оценивающих ответы тестируемых (0,84–0,96). Шкала оценки была также пятиуровневая, как было указано выше. Оценивались лексика, грамматика, акцент (произношение), беглость речи и понимание. В 1984 г. эти компоненты были изменены, и оценивались понимание, лексикализация (семантический аспект), грамматическая точность, беглость и дискурсивная компетенция.

д) Коммуникативное направление, представленное моделями М. Кэнела и М. Свайна [92], а также Л. Бахмана [87]. Это направление основывает оценку коммуникативной компетенции на определении степени развитости компонентов КК. К этому направлению относится и видоизмененный подход, принятый в 1982 г. ACTFL (Американским Советом преподавателей иностранных языков), которому дало название ориентированное на общее владение языком (proficiency-oriented) обучение и тестирование.

Оценка общего уровня владения языком производится по трем компонентам – коммуникативные функции (functions), тематика (topics, content/context) и точность (приемлемость, качество и точность выражений).

Характерными и положительными чертами этого периода являются стремление к большей коммуникативности в тестировании, к большей контекстуальности, аутентичности (real-lifeness), к большему вниманию к преподавательским и диагностическим тестам.

В перспективе можно ожидать усовершенствование коммуникативно-ориентированного тестирования с широким диапазоном коммуникативности тестов в зависимости от потребностей учащихся (программ, учебников) и уровня их владения ИЯ. Все больше внимания, видимо, будет обращено на максимально полное использование возможностей ЭВМ в целях индивидуализированного адаптивного, интерактивного и самостоятельного (для самоконтроля) тестирования.

4.2. РАЗВИТИЕ ЛИНГВОДИДАКТИЧЕСКОГО ТЕСТИРОВАНИЯ В СССР

В развитии отечественного тестирования можно выделить два периода – довоенный (1920–1936 гг.) и послевоенный – с середины 60-х годов.

В начале 20-х годов вместе с созданием в СССР новой системы образования педагоги решили по-новому подойти к проблемам профотбора и контроля успеваемости. Наряду с психологическими тестами общих и специальных способностей стали разрабатываться и довольно широко применяться педагогические тесты. Для этого педагоги использовали как зарубежную теорию и практику тестирования, так и отечественные разработки.

Однако, вскоре после первых успехов появилась переоценка и абсолютизация возможностей психологических и педагогических тестов. Теория и практика тестирования еще недостаточно установились, и созданные тесты были некачественными, допускались серьезные ошибки в интерпретации их результатов.

Возобновление использования метода тестов в педагогике в начале 60-х годов связано с применением программированного обучения и программированного контроля. (См., например, статьи С.И. Воскерчьяна [8] и Т.А. Ильиной [20].) Во второй половине 60-х годов появились обзорные статьи Л.В. Банкевича [3], И.Н. Гальцева [10], Г.Ю. Князевой [23], Т.А. Королевой [39], С.П. Суворова [72], а также брошюры И.А. Цатуровой [77] и С.П. Макушевой [51] об опыте разработки языковых тестов за рубежом. Были защищены первые диссертации по теории и практике языковых тестов – И.А. Цатуровой [79] и И.А. Рапопортом [60]. Всего до 1987 г. защищено около 20 кандидатских диссертаций по тестированию в области ПИЯ.

В последнее десятилетие появились пособия И.А. Рапопорта и др. [62], Л.В. Банкевича [4] и Э.А. Штульмана [82], в которых наиболее полно и систематизированно освещаются вопросы теории и практики ЛДТ.

Уже в начале 70-х годов начали образовываться коллективы и центры, занимающиеся разработкой и опробованием ЛДТ, а также их теорией. Наиболее плодотворными среди них можно назвать кафедру ИЯ Таганрогского радиотехнического института (ТРТИ) во главе с И.А. Цатуровой, работающих вместе с учеными УССР и ЭССР И.А. Рапопорта, И. Соттер и Р. Сельг, преподавателей МГИИЯ им. М. Тореза проф. С.К. Фоломкину, Н.С. Шебеко и др., коллектив кафедры ИЯ Ленинградского политехнического института во главе с проф. Л.В. Банкевичем, коллектив кафедры ИЯ Таллиннского политехнического института во главе с автором настоящего пособия и многие другие коллективы и ученых. Осветим вкратце деятельность некоторых из них.

Под руководством И.А. Цатуровой в ТРТИ с 1969 по 1971 г. была составлена и опробована система локально стандартизованных тестов учебных достижений для студентов I–VIII семестров по трем иностранным языкам, а также тест исходного (стартового) уровня. С 1974 по 1976 г. совместно с учеными из других вузов был составлен и опробован всесоюзный стандартизованный тест по грамматике, а затем и по лексике для студентов II курса [54]. В ТРТИ организован факультет повышения квалификации по тестированию для преподавателей из других вузов и проводятся регулярные школы-семинары с представительным форумом тестологов страны. Во многих статьях авторов ТРТИ освещены самые актуальные вопросы теории и практики ЛДТ.

В 1972 г. в 14 вузах страны опробовался комплексный тест, составленный группой ученых из МГИИЯ им. М. Тореза под руководством Ю.А. Веденяпина. Целью тестирования являлось выявление исходного уровня обучаемых для разработки программированных пособий (обуча-

ющих программ) по трем иностранным языкам [6]. Результаты этих тестов приведены в статье А.Г. Решетова [66], а их сопоставительный анализ – в статье В.А. Коккоты [25].

Разнообразными являются направления исследований И.А. Рапопорта. Это обучаемость [57; 13], лингвистические способности и память [59], стандартизованные тесты, опробованные в г. Николаеве и ЭССР совместно с Р. Сельг и И. Соттер [62; 63; 64], вопросы методологии составления языковых тестов, совершенствование техники тестирования, вопросы надежности и валидности тестов [58] и др. Библиография статей И.А. Рапопорта, Р. Сельг и И. Соттер приведена в их пособии [62].

Значительным является вклад в отечественное лингводидактическое тестирование проф. С.К. Фоломкиной (статьи по методике, лекции по тестированию на ФПК при МГИИЯ им. М. Тореза, руководство диссертантами и т.п.), а также коллектива ЦКМОИЯ, особенно Н.С. Шебеко [45; 46; 65; 70] при освещении теоретических проблем контроля в обучении ИЯ.

В Таллинском политехническом институте (ТПИ) в период 1974–1978 г. был составлен и опробован локально стандартизованный тест (АТ-78), который был дважды опробован в вузах ЭССР, а также в Рижском ПИ и Каунасском ПИ [29; 36; 37]. Большое внимание в ТПИ уделяется вопросам техники тестирования [9; 34; 32], в частности тестам типа клоуз [37; 26; 31], а также тестам аудирования и лингвистических способностей. разработан локально стандартизованный тест способностей АТ-80 с выборочными ответами [27; 35]. С 1982 г. идет работа по составлению системы вузовских стандартизованных тестов по английскому языку, из которых опробование прошли тест исходного (стартового) уровня (ИТА-84) и тест для студентов II курса неязыковых вузов (АТУ-88). В опробование тестов ИТА-82 и -83 участвовало 75 вузов девяти союзных республик с привлечением примерно 7,5 тыс. студентов ежегодно, а в опробовании тестов АТУ-85 и -86 – 35 вузов девяти союзных республик с привлечением примерно 4,2 тыс. студентов [31]. Тест АТУЛЕКС-86 и -88 составлялся совместно с тестологами Каунасского ПИ. ТПИ провел три региональных семинара с изданием сборников тезисов докладов на иностранных языках.

В целом можно констатировать, что отечественное тестирование давно вошло в структурно-психометрическую стадию по периодизации Б. Спольского, но во многих случаях к качеству разрабатываемых тестов предъявляется недостаточная требовательность, не анализируют трудность и дифференцирующую способность заданий теста. Несомненное негативное влияние оказало отсутствие в 70-х годах отечественных пособий по проблемам теории и практики тестирования.

5. ОБЩАЯ МЕТОДИКА РАЗРАБОТКИ ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

5.1. ПЛАНИРОВАНИЕ ЛИНГВОДИДАКТИЧЕСКОГО ТЕСТИРОВАНИЯ

Стадии и этапы тестирования были приведены выше на рис. 1. Планирование тестирования охватывает: 1) планирование самого ЛДТ как центрального звена; 2) планирование отдельных этапов тестирования (уточнение объемов работы) и 3) составление графика работ. Планирование тестирования во многом зависит от того, разрабатывается ли новый тест, дополняется/перерабатывается ли имеющийся или будет использован готовый тест.

В специальной литературе планирование тестирования освещено весьма поверхностно, только у Д. Харриса [116] можно найти описание планирования некоторых этапов языковых тестов. Видимо предполагается, что при тестировании применимы общие принципы планирования, при которых учитываются наличные ресурсы, объемы работы, производительность труда, сроки и т.п. Однако ЛДТ имеют свои специфические особенности.

Планирование лингводидактического тестирования состоит из следующих действий:

1. Определение общих и конкретных целей и функций составляемого ЛДТ, т.е. будет ли тест составляться для контроля текущей, промежуточной или итоговой успеваемости, для установления исходного уровня обученности или для исследовательских целей; начинается или дополняется существующий банк данных (банк заданий, текстов, тестов и т.п.) и информационная система кафедры или вуза, например, учета успеваемости, учета нагрузки преподавателя и студента и многое другое. Такие банки и системы можно вначале составить в виде картотеки с учетом ее перевода позже на ЭВМ.

2. Определение общего характера теста – языковой, речевой или комбинированный и определение подлежащих проверке обобщенных объектов тестирования (лексика, грамматика, чтение и др.). Ниже в примерах приведены только комплексные тесты, состоящие из нескольких частей, например, по лексике, грамматике, чтению и аудированию. В случае ПТ и ЛСТ возможно, что с точки зрения внесения изменений будет более целесообразно составление одними преподавателями теста (части теста) по лексике, а другими – по грамматике и т.д.

3. Определение ресурсов рабочей силы и времени – на составление, проведение (аудиторное время), обработку данных для уточнения степени стандартизации составляемого теста. Желательно стремиться шаг за шагом увеличивать степень стандартизации: применять в традиционной контрольной работе методы обработки тестов с целью анализа пригодности их заданий для тестирования, использовать эти же задания для

преподавательских тестов, вычисляя при возможности, кроме трудности заданий, их дифференцирующую способность. Наиболее приемлемые задания можно использовать для локально стандартизованного теста вуза/школы, города (республики). На выполнение всего указанного выше требуются дополнительное время и силы. В таблице 12 приведены возможные сочетания целей и степеней стандартизации ЛДТ. (См. также табл. 5.)

Таблица 12

Приемлемые сочетания целей и характера ЛДТ

Цели ЛДТ: для проверки успеваемости	Функции ЛДТ			Степень стандартизации		
	диагностические	учебные достижения	общее владение ИЯ	СТ	ЛСТ	ПТ
текущей	+ +	+ +	+	–	+	+ +
промежуточной	+	+ +	+ +	+	+ +	+
итоговой	+	+ +	+ +	+ +	+ +	+
степень коммуникативности	низкая	средняя	средняя/высокая	средняя/высокая	средняя	средняя/низкая

Примечание: «+ +» – составление весьма целесообразно; «+» – целесообразно, но имеются недостатки; «-» – нецелесообразно.

4. Уточнение соотношения составляемого ЛДТ с применяемой методикой обучения, программными целями и задачами рабочей программы, а также с существующей системой контроля для определения степени коммуникативности теста. Желательно стремиться к большей степени коммуникативности, но следует учесть, что с ростом коммуникативности теста растут и расходы времени и сил на составление теста, а при высококоммуникативных (субъективных) тестах – и на проведение и обработку результатов ЛДТ.

5. Анализ рабочей программы, учебных материалов (учебника, текстов, лексических минимумов) для определения объектов тестирования, объема тестируемого языкового материала и перечня проверяемых навыков и умений. В случае преподавательских тестов объем анализируемого материала более конкретный и ограниченный, но и времени на проведение теста меньше. Сложнее отбор языкового материала и тестируемых навыков и умений в случае ЛСТ и тестов промежуточной и итоговой успеваемости, так как критерии отбора далеки от совершенства. При планировании количества заданий необходимо учесть запас 25–40 %, поскольку опробование непременно выявит какую-то долю неудачных,

слишком легких или слишком трудных заданий. Если требуемое количество заданий слишком большое для данного времени проведения ЛДТ, то нужно увеличивать количество форм теста. Для окончательного теста необходимо планировать минимум две равнотрудные формы, чтобы исключить «сотрудничество» сидящих рядом тестируемых.

6. Выбор техники тестирования в соответствии с объектами тестирования, результатами анализа. На этом этапе следует решить окончательно степень коммуникативности заданий теста (их контекстуальность, аутентичность). Целесообразно продумать возможность использования заданий, частей или целых готовых ЛДТ, которые подходят для рассматриваемых целей и условий. Например, можно использовать данные из имеющегося в Центральном кабинете методики обучения (МГИИЯ) отчета по разработке всесоюзных тестов серии ИТА. В этом отчете ТПИ [30] приведены тексты всех заданий, их трудность в виде процента правильных ответов, типичные ошибки в ответах на 385 заданий по лексике при их внеконтекстной семантизации. Такие же данные имеются там и о тестах серии АТУ.

7. Уточнение состава разрабатываемого теста – количества заданий в частях, методическое и тесто-экономическое обоснование проекта теста и перепланирование теста с учетом данных п. 2, если невозможно уложиться в заданное для проведения теста время. Могут быть изменены, например, техника тестирования, соотношение количества заданий в разных частях теста, увеличено количество форм теста. При оценке времени на проведение теста следует учесть, что 5–10 минут требуется для организации проведения теста – раздачи и сбора материалов и т.п.

5.2. СОСТАВЛЕНИЕ ЗАДАНИЙ ЛИНГВОДИДАКТИЧЕСКИХ ТЕСТОВ

Под заданием ЛДТ понимается любой стимул, предназначенный для возбуждения измеряемых речевых действий тестируемого, позволяющих оценить ее/его лингвистическую и/или коммуникативную компетенцию. Каждая группа (часть, субтест) однотипных заданий обычно имеет заголовок, инструкцию и пример выполнения заданий.

Требования к заданиям ЛДТ. Задания языковых тестов должны отвечать следующим основным требованиям:

1) соответствие формы и характера заданий теста целям и объектам тестирования, а также пройденному языковому материалу для обеспечения максимально возможной валидности теста;

2) доступность по форме и содержанию, а также посильность по трудности заданий для данного контингента тестируемых (трудность заданий в пределах от 30–35 % до 80–85 %);

3) уверенная и стабильная дифференциация более сильной и более слабой частей контингента по результату данного задания (ДС более 0,4) для обеспечения надежности теста;

4) соблюдение в языковых тестах следующего принципа: в одном задании только одна трудность – та, которая подлежит проверке в данном задании;

5) наличие обучающей ценности, ситуативности и проблемности задания, основанного на интересном содержании и представленного на должном идейно-воспитательном уровне для обеспечения достаточной мотивации и нужных установок тестируемых при выполнении заданий;

6) учет оптимальной средней длины заданий (9 ± 1 слов без пропуска) в основе (тексте) задания;

7) правильное соотношение лингвистической и экстралингвистической ситуации в задании теста;

8) соответствие лингвистического уровня заданий теста уровню усвоения языкового материала; рецептивное и репродуктивное владение контролируется в основном на уровне фонем, слов, словосочетаний и предложений;

9) наличие сбалансированных отвлекающих ответов при использовании техники множественного выбора (ТМВ).

Требования к заданиям речевых тестов еще недостаточно разработаны. Так, В.Н. Шацких и Н.М. Липатникова [65] приводят следующие требования к контролю (приводятся сокращенно): 1) нацеленность на речевое умение, а не на знания о языке; 2) одноязычный, устный, творческий и скрытый характер; 3) интересное содержание и привлекательная форма выполнения, нацеленность больше на оценку, а не на отметку и др.

Главное требование к заданиям речевых тестов – это обеспечение их максимально возможной степени коммуникативности (по положениям таблиц 3 и 4). Из них важнейшие – непосредственный, интерактивный характер речевых заданий, направляемых на контекстуальные значения высказываний в аутентичных ситуациях общения. Важными представляются еще требования охвата максимального количества компонентов коммуникативной компетенции и широкого использования аутентичного языка. Дополнительно к указанным отметим еще требования функциональности (соответствие характера и содержания заданий потребностям и интересам тестируемого) и познавательной ценности (новизна, творческий характер) тематики и выполняемых речевых действий в заданиях речевых тестов.

Классификация заданий ЛДТ. Задания ЛДТ весьма разнообразны, и для их классификации можно исходить из следующих признаков:

1) форма предъявления стимула – наглядная (динамическое или статическое изображение) или вербальная – внеконтекстная или контекстная (микро- и макроконтекстная);

2) форма представления ответа – наглядная, цифро-буквенная, вербальная – внеконтекстная или контекстная;

3) характер ответа по подсчету результатов – объективный (по ТМВ), полуобъективный (короткий однозначный ответ) или субъективный (со СКО);

4) модальность стимула и ответа (могут не совпадать) – устная или письменная;

5) используемый в стимуле и в ответе язык (могут не совпадать) – родной или изучаемый иностранный язык.

Из приведенных признаков следует, что их возможных сочетаний сотни, однако, исходя из методической целесообразности и приведенных в таблицах 13–16 данных, число приемлемых сочетаний ограничивается несколькими десятками.

Таблица 13

Оценка возможных сочетаний форм стимула и ответа в заданиях ЛДТ

Формы представления (формирования) ответа	Формы предъявления стимула				
	наглядный (статическое или динамическое изображение)	одно слово (словосочетание)	группа слов (часть предложения)	контекст с пропуском	контекст без пропуска
Наглядная (картинки и т.п.)	–	+	+	–	+
Цифро-буквенная	+	+ +	+ +	+ +	+
Одно слово (словосочетание)	+	+	+	+ +	+ +
Группа слов, изолированных от контекста	+	+ +	+ +	+ +	+
Контекстный ответ	+ +	–	+	–	+ +

Примечание: «+ +» – наиболее приемлемое сочетание; «+» – сочетание с недостатками; «–» – невозможное или нецелесообразное сочетание.

Основные типы заданий в основном заимствованы у Л.В. Банкевича [4]. Их применимость в языковых и речевых тестах показана в таблице 14. Приведем их краткое описание:

1. Завершение (чаще заполнение пропуска в предложении или в тексте – тест типа клоуз) является одним из наиболее распространенных типов заданий по ТМВ. Этот тип задания используется в основном для

проверки лексико-грамматических навыков чтения и аудирования. В случае теста типа клоуз по ТМВ этот тип заданий измеряет в первую очередь навыки и умения чтения.

Стимул задания этого типа называется основой, или текстом. Методика составления основы была описана в разделе экономичности тестов. Было указано, что текст задания может быть написан самим составителем, его можно выбрать среди пройденного языкового материала или среди непройденного, но аналогичного (такого же по уровню), а также взять из аутентичного материала (книги, газеты, журналы, учебники, энциклопедии, словари и т.п.).

Стимул лексического задания может создать такую тестовую ситуацию, которая подсказывает слово, которым нужно завершить основу (заполнить пропуск). Однако довольно широко применяются и такие задания, в которых только сопоставление основы с вариантами ответов позволяет выбрать среди них правильный. Р. Ладо [127] считает последнюю возможность лучшей.

Таблица 14

Оценка применимости основных типов заданий в языковых и речевых тестах

Основные типы заданий и ответов	Языковые тесты			Речевые тесты		
	наглядный стимул	внеконтекстные стимулы	микроконтекстные стимулы	чтение	аудирование	говорение
I. Избирательные ответы 1. Завершение (заполнение) пропусков	–	+	+ +	+ +	+ +	–
2. Синоним (парафраза)	–	–	+ +	+	+	–
3. Двойной (альтернативный) выбор (правильно – неправильно)	+	+ +	+ +	+	+	+
4. Группировки	–	+ +	+	–	–	–
5. Сопоставление (сличение)	+	+ +	+	–	–	–
6. Интерпретация	+	+	+ +	+ +	+ +	+ +

Продолжение

Основные типы заданий и ответов	Языковые тесты			Речевые тесты		
	наглядный стимул	внеконтекстные стимулы	микроконтекстные стимулы	чтение	аудирование	говорение
II. Свободно конструируемые ответы 1. Свободный короткий ответ	+ +	+	+	+	+	+
2. Свободный полуоткрытый ответ (завершение)	–	–	+ +	+ +	+ +	–
3. Ответ на вопрос или реакция на стимул (трансформация и др.)	+ +	+	+	+ +	+ +	+ +
4. Интерпретация	–	+	+ +	+	+	+ +
5. Парафраза, трансформация	–	+	+ +	+	+	+
6. Исправление	–	–	+	+ +	–	–

Примечание: «+ +» – наиболее приемлемое сочетание; «+» – сочетание с недостатками; «–» – невозможное или нецелесообразное сочетание.

Рассмотрим подробнее вопрос о характере отвлекающих ответов (см. также [4; 127; 99; 166; 116; 112]). Х. Гудрич [112] провел экспериментальное исследование привлекательности (процент выбора тестируемыми) и дифференцирующей способности следующих типов отвлекающих (по порядку эффективности): 1) контекстуальная отвлекающая, 2) антоним правильного ответа, 3) неправильный синоним

(в данном контексте), 4) аффиксальная отвлекающая (к правильному ответу прибавлен или отнят от него префикс или суффикс), 5) фонетико-графически интерферирующая с родным (арабским у Х. Гудрича) отвлекающая, 6) графическая отвлекающая (одна-две буквы прибавлены к правильному ответу или отняты от него), 7) клоуз-отвлекающая, полученная по предварительным клоуз-тестам (синтактико-семантическая отвлекающая), 8) случайное слово.

Р. Ладо [127] советует, чтобы один вариант ответа был отвлекающим по форме (напоминал бы форму правильного ответа), другой вариант ответа – отвлекающим по контексту и третий вариант ответа – отвлекающим по значению (многозначным словом). Отвлекающие должны иметь примерно одинаковые длину и трудность, а также область знаний с правильным вариантом ответов.

При отборе отвлекающих необходимо учесть еще то, что все отвлекающие и правильный ответ должны содержать одну и ту же часть речи, поскольку все варианты ответов взаимосвязаны тем, что тестируемый сравнивает их в ходе поиска правильного ответа. Важно учесть еще оптимальность семантического расстояния [4] между вариантами ответов, т.е. они должны быть в равной степени детальными и обобщенными, как отмечает Дж. Кларк [98]. Чтобы исключить «ловушки» для более сильных тестируемых, лучше искать отвлекающие с достаточно большим семантическим расстоянием от правильного ответа. Как следует из исследования Х. Гудрича, далекая отвлекающая (максимальное семантическое расстояние), используемая для слепого угадывания, наименее эффективна и ее следует использовать в последнюю очередь.

2. Синоним (парафраза) по ТМВ является микроконтекстным лексическим заданием, в котором тестируемому слову (обычно подчеркнуто в предложении) следует найти синоним или парафразу среди вариантов ответов. Этот тип задания более естественный, чем завершение, однако имеет два недостатка. Во-первых, неизвестно, какого слова тестируемый не знает, если ошибается, и во-вторых, часто невозможно найти подходящий синоним или парафразу среди пройденного языкового материала.

Типы заданий 1 и 2 используются в частях лексики тестов ТОЭФЛ, ИЭА, Мичиганский тест и др.

3. Двойной (альтернативный) выбор является модификацией лексических заданий по ТМВ. Для этого типа задания характерно сравнительно большое значение смыслового содержания основы-текста задания. При внеконтекстной семантизации используется сравнение синонимичных или антонимичных пар слов (ответы типа «одинаковое – различное значение»). В случае микроконтекстных заданий используются ответы типа «правильно – неправильно». Недостатком этого типа является большая вероятность случайного угадывания.

4. Группировками называются внеконтекстные лексические задания по ТМВ, в которых правильный вариант ответа входит в группу слов, являющуюся заданием. Л.В. Банкевич [4] приводит три основных дей-

ствия с группировками: 1) исключение того члена группировки, который обозначает нечто другое, чем все остальные члены; 2) обобщение группировки, т.е. обобщение смысловых отношений между словами группировки и нахождение родового слова; 3) расположение по порядку, при котором за основу принимается какой-либо признак естественного ряда, например, название дней недели и т.п.

5. Сопоставление (сличение) лексических единиц (ЛЕ) используется согласно таблице 15 в довольно многих разновидностях. При использовании ТМВ каждое значение (ЛЕ) может иметь 1) свои 3–5 сопоставляемых вариантов ответов или 2) общие с другими заданиями варианты ответов (сопоставление списков или столбцов ЛЕ на родном или иностранном языках). Это самый экономичный тип заданий, имея в виду этапы составления, проведения и обработки результатов теста.

Для уменьшения времени на механическое просматривание – поиск правильного ответа в длинном списке-столбце ответов – следует либо ограничиться 7–8 вариантами ответов (на 4–5 заданий), либо привести варианты ответов в алфавитном порядке, известив тестируемых об этом в инструкции [34].

Таблица 15

Модификации тестового задания типа «сопоставление»

Форма представления (формирования) ответа	Форма стимула задания			
	Внеконтекстная		Микроконтекстная	
	Наглядная	ИЯ	РоЯ	ИЯ
1. Наглядная	–	+	–	+
2. Внеконтекстная ИЯ	+ СКО, ТМВ	+ + СКО, ТМВ	+ СКО, ТМВ	+ + СКО, ТМВ
РоЯ	–	+ СКО, ТМВ	–	+ СКО, ТМВ
3. Микроконтекстная ИЯ	+ + СКО	+ + ТМВ	+ СКО	+ СКО, ТМВ
РоЯ	–	+ ТМВ	–	+ СКО, ТМВ

Примечание: см. под таблицей 14.

6. Интерпретация по ТМВ или со СКО является наиболее коммуникативным типом заданий языкового теста и заключается в сопоставлении частей (слов, словосочетаний, предложений) речевого акта, при котором

следует с помощью интерпретации и понимания смысла высказываний соединить разделенные в задании и вариантах ответов части. Этот тип заданий можно использовать на более продвинутых этапах обучения.

Рассмотрим вкратце также основные типы заданий тестов с субъективной оценкой ответов. Первые два типа заданий – с коротким (1–2 слова) ответом и со свободным полуоткрытым ответом (завершение ответа-высказывания 1–2 словами) – можно условно назвать полуобъективными в том случае, если эти задания составлены так, что ответы достаточно однозначны. На начальных и средних этапах обучения словарный запас тестируемых относительно невелик и репертуар синонимичных значений ограничен, поэтому здесь можно использовать механическую проверку ответов путем сопоставления правильного ответа на трафарете с ответами тестируемых на ответном листе.

1. Задание со свободным коротким ответом хорошо подходит для названия предметов, действий и т.п. на картинах при проверке синонимичных или антонимичных значений. Полуобъективным можно считать также свободные короткие (однословные) ответы на задания тестов типа клоуз, которые могут быть сопоставлены с правильными ответами по тексту этого теста, т.е. контекст в значительной степени диктует ответ тестируемого.

2. Задание со свободным полуоткрытым ответом удобно использовать для проверки поискового и ознакомительного чтения. Тогда тестируемому предлагается сформулировать основную информацию (о чем речь в таком-то разделе, что имеется в виду), а не дается ряд выборочных ответов, для которых иногда довольно трудно подобрать эффективные отвлекающие ответы. То же самое относится к некоторым объектам тестирования аудирования. Обычно таких заданий не более 5–10, и проверка коротких предложений или словосочетаний даже у всей учебной группы не занимает слишком много времени, а сами задания сравнительно легко составляются.

3. Задания, требующие свободного ответа на вопрос или реакцию на стимул-высказывание, обычно не относят к тестовым заданиям, и они применяются только в тестах устной речи с четкой процедурой проведения и оценки ответов тестируемых. Однако из-за устной формы ответов и их разнообразного характера такие ответы трудно оценить. Несколько легче оценить короткие письменные ответы на стимулы (вопросы, утверждения, высказывания) о содержании прочитанного и услышанного, а также со зрительной опорой на картину, диафильм и т.п.

4. Задания на интерпретацию со СКО имеют примерно аналогичный характер с заданиями по ТМВ и занимают промежуточное место среди типов заданий 2 и 3 со СКО.

5. Задание парафразы со СКО сводится к письменному или устному объяснению значения слов, словосочетаний или предложений. Аналогичным является тип заданий, именуемый трансформацией, который используется в грамматических тестах в виде заданий типа: преобразовать настоящее время в прошедшее, изъявительное наклонение в сослагательное.

6. Задания на исправление используются в тесте ТОЭФЛ в виде редакторского теста, в котором в тексте нужно найти не подходящие по смыслу или ошибочные слова (в случае ТМВ большинство слов пронумерованы).

Составление внеконтекстных и микроконтекстных заданий ЛДТ рекомендуется производить на карточках величиной в половину страницы, поместив на эти карточки: 1) данные о тесте и тестируемом объекте (повторяемость ЛЕ или грамматической формы, их частотность в языке, трудность по другим тестам в других условиях и т.п.); 2) текст одного или нескольких заданий и использованные (неиспользованные) варианты ответов; 3) данные об опробовании данного задания в предтесте, а затем и в окончательном тесте (ИТ, ДС и другие данные).

5.3. ОБРАБОТКА И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ ТЕСТА

При обработке результатов предтеста основное внимание уделяется анализу заданий для составления качественного окончательного теста. В случае второго опробования предтеста главным является выявление качества (надежности, валидности и других показателей) теста, а в случае проведения/использования окончательного/готового теста самыми важными являются результаты теста.

Целью обработки результатов теста в минимальном объеме является: 1) получение средних данных о результатах теста всего контингента тестируемых, которые характеризуют эффективность (успешность) обучения и обученность всего контингента тестируемых и 2) получение данных о результатах каждого отдельного тестируемого, которые характеризуют внутригрупповые различия в обученности. При максимальном объеме обработки к указанному прибавляется анализ заданий – индекса трудности (ИТ), дифференцирующей способности (ДС) и эффективности отвлекающих, вычисление надежности и, при возможности, валидности и других показателей качества теста.

В настоящее время распространяются приемы полной обработки результатов языковых тестов на ЭВМ (в ТПИ с 1985 г. на ЭВМ переведена обработка тестов серии АТУ), но ниже рассматриваются приемы обработки результатов теста вручную, которые могут пригодиться при обработке результатов преподавательских тестов.

Важным принципом, позволяющим сэкономить время на проверку ответных листов, является работа с неправильными ответами на первых этапах обработки. Это означает, что с помощью трафаретов отмечаются места неправильных ответов и сразу же производится подсчет их количества. В случае ответов по ТМВ через окна трафаретов кружками или галочками отмечают недостающие правильные ответы, а при СКО трафарет располагают слева близко к колонке ответов тестируемого и сравнивают слова на трафарете и на ответном листе. Если нужно вычислить количество правильных ответов, то из количества заданий в тесте (в части теста) вычитают количество неправильных ответов. Кроме

выигрыша во времени работа с неправильными ответами позволяет выявить типичные ошибки у данного контингента, анализ которых способствует более правильной организации учебного процесса.

Предлагается следующая последовательность действий при обработке результатов ЛДТ:

1. Проверка ответных листов, отметка на них мест неправильных ответов и их количества в каждой части теста.

2. Перенос индивидуальных результатов частей тестов с ответных листов на сводные таблицы и вычисление средних результатов по нижеследующим формулам. Выражение среднего с помощью процентной доли позволяет сравнить средние результаты разных тестов и перейти на пятибалльную шкалу. Индивидуальный средний результат теста (его части) по неправильным ответам M_0 равен в процентах:

$$M_0 = \frac{KO_x}{n} \times 100, \qquad (2)$$

где KO_x – количество ошибочных (неправильных) ответов данного тестируемого и n – количество заданий в данной части (тесте);

Поскольку количество правильных ответов $K\Pi_x$ равно:

$$K\Pi_x = n - KO_x, \qquad (3)$$

то средний результат тестируемого в процентах через правильные ответы равен:

$$M_n = 100 - M_0 \qquad (4)$$

или

$$M_n = \frac{K\Pi_x}{n} \times 100. \qquad (5)$$

Хотя коррекция индивидуальных средних результатов теста на случайное угадывание является недостаточно обоснованной по мнению Дж. Хитона [119] и Д. Харриса [116], все же приведем соответствующую формулу:

$$M_к = K\Pi_x - \frac{KO_x}{K - 1}, \qquad (6)$$

где К является количеством неправильных (отвлекающих) вариантов ответов.

Средний результат всего контингента тестируемых (средняя трудность теста) М равен:

$$M = \frac{\sum_1^N M_n}{N} \quad или \quad M = \frac{\sum_1^N K\Pi_x}{N \cdot n}, \qquad (7)$$

где N – количество тестируемых, выполнивших данный тест с общим количеством правильных ответов $\sum_{1}^{N} КП_x$.

3. Анализ заданий может быть произведен прямо на ответном листе, на аналогичной ответному листу таблице или на карточках заданий. Хотя это кажется более трудоемким, опыт показывает, что целесообразнее сразу выявить как трудности заданий, так и их дифференцирующую способность (ДС) и эффективность дистракторов.

Основным принципом анализа задания в таком случае является то, что указание или знак того, что данный тестируемый ошибся в определенном задании, в виде, например, черточки или точки переносится в таблицу для анализа не просто на место этого задания, а используя «адрес» ошибочного ответа, т.е. ту отвлекающую, которую этот тестируемый выбрал, этот знак (черточка) проставляется на соответствующее место отвлекающего ответа в таблице. Другими словами, производится точный перенос ошибочных выборов тестируемого. Лучше всего в качестве такой таблицы использовать ответный лист этого же теста или его увеличенную аналогию.

Поскольку ДС выражается через разницу между результатами (ошибками) слабой и сильной частей контингента, то ответные листы нужно сначала расположить по порядку убывания количества неправильных ответов в тестах (его частях), тогда слабая часть образуется отбором 27–33 % верхних ответных листов, а сильная часть – таким же количеством нижних ответных листов. При количествах тестируемых около 25–35 указанная доля составляет 33 %, при количестве 40–60 – 30 %, а более 75 тестируемых – 27 %.

Для каждой части контингента берут отдельный ответный лист или выделяют часть таблицы для переноса на них знаков (черточек) ошибок тестируемых, поскольку номера анализируемых заданий одинаковы для всех частей контингента. Когда перенос неправильных ответов закончен во всех трех частях контингента, то индекс трудности (ИТ) в процентах данного задания вычисляется по следующей формуле:

$$ИТ = \frac{КО_{си} + КО_{сл} + КО_{ср}}{N_{си} + N_{сл} + N_{ср}} \times 100, \qquad (8)$$

где $КО_{си}$ – количество неправильных ответов в сильной части, $КО_{сл}$ – в слабой части и $КО_{ср}$ – в средней части контингента, а $N_{си}$, $N_{сл}$, $N_{ср}$ – количество тестируемых соответственно в сильной, слабой и средней частях контингента тестируемых.

Трудность задания по правильным ответам в процентах можно найти по формуле:

$$R = 100 - ИТ. \qquad (9)$$

ДС данного задания определяется с помощью следующих формул:

$$ДС = \frac{КО_{сл} - КО_{си}}{N_{сл}} \quad \text{или} \quad ДС = \frac{КП_{си} - КП_{сл}}{N_{си}}. \qquad (10)$$

Аналогично можно найти эффективность (привлекательность) отвлекающих, т.е. сколько тестируемых (процентуально) выбрало тот или иной отвлекающий вариант ответа, и насколько хорошо отвлекающие разделяют сильных и слабых тестируемых ($ДС_0$).

4. Вычисление показателей качества теста целесообразно начать с вычисления среднеквадратичного отклонения среднего результата теста, которое производится по формуле:

$$\sigma_t = \sqrt{\frac{\Sigma(КП_x - M)^2}{N}} \quad \text{или} \quad \sigma_t = \sqrt{\frac{\Sigma(КП_x)^2}{N} - M^2}, \qquad (11)$$

где $КП_x$ – количество правильных ответов данного тестируемого, M – среднее количество правильных ответов всех тестируемых, N – количество тестируемых.

Вычисление σ_t удобно производить с помощью таблицы – списка тестируемых с их результатами по тесту. По результатам каждого тестируемого ($КП_x$) вычисляются разности КП – М, которые затем возводят в квадрат и суммируют ниже. Полученная сумма делится на N и берется квадратный корень. Вычисления σ_t удобно произвести на калькуляторе, накопляя в его памяти суммы квадратов указанных разниц или $КП_x$ (по второй формуле). В таком случае таблица не требуется.

Для сравнения с другими тестами удобнее представить среднеквадратичное отклонение в виде коэффициента вариации в процентах:

$$V = \frac{\sigma_t}{M} \times 100. \qquad (12)$$

Надежность теста можно подсчитать по разным формулам. Приведем некоторые из них, наиболее удобные для преподавательских тестов. Если известно σ_t всего теста легко можно определить еще и среднеквадратичные отклонения по четным и нечетным заданиям (σ_1 и σ_2 соответственно), то надежность теста выражается по формуле Л. Гутмана через дисперсию (σ^2):

$$r_t = 2\left(1 - \frac{\sigma_1^2 - \sigma_2^2}{\sigma_t^2}\right). \qquad (13)$$

Удобной является также один из вариантов формулы Кюдера-Ричардсона, сокращенно формулы KR = 21:

$$r_t = \frac{n}{n-1}\left(1 - \frac{M(n-M)}{n \cdot \sigma_t}\right), \qquad (14)$$

где M – средний непроцентный результат теста и n количество заданий в тесте. Поправка $\frac{n}{n-1}$ нужна при количестве заданий до 70–80.

Более точные результаты дает формула KR – 20 в виде:

$$r_t = \frac{n}{n-1} \cdot \frac{\sigma_t^2 - (\Sigma p - \Sigma p^2)}{\sigma_t^2}, \tag{15}$$

где p – трудность задания без процента, т.е.

$$p = \frac{\Sigma КП_x}{N}. \tag{16}$$

В формуле KR – 20 суммируются непроцентные трудности всех заданий, и из этой суммы вычитается сумма квадратов указанных трудностей заданий.

Для вычисления валидности теста имеется также несколько способов. Иногда индивидуальным результатам тестов, по которым определяется сопоставительная валидность одного из них, удобно присваивать порядковые номера (ранги), по которым тестируемые распределяются соответственно своим результатам по тесту среди всего контингента тестируемых. Это особенно целесообразно при коррелировании (валидации) результатов теста с оценками успеваемости (средними за месяц, семестр и т.п.). Соответствующая формула ранговой или порядковой корреляции Спирмена имеет вид:

$$\rho = 1 - \frac{6 \cdot D^2}{N(N^2 - 1)}, \tag{17}$$

где D – разность рангов (порядковых номеров или мест тестируемых в списке по результатам теста), а N – количество тестируемых. Таким образом, в списке тестируемых их результатам по тесту сначала присваивают порядковые номера (в отдельной графе), т.е. находят, например, наименьшее количество неправильных ответов (это № 1), затем следующий результат (это № 2) и так далее до конца списка. Такая же процедура проводится по результатам другого теста или по баллам успеваемости. Квадраты разниц этих порядковых номеров (рангов) удобно вычислять на микрокалькуляторе и накоплять в его памяти для последующих расчетов.

Если тестируемые по своим результатам теста имеют так называемое нормальное распределение, то для определения валидности теста правомерно использование коэффициента корреляции K по К. Пирсону:

$$K = \frac{\Sigma (КП_1 - M_1)(КП_2 - M_2)}{N \cdot \sigma_1 \cdot \sigma_2}, \tag{18}$$

где $КП_1$ и $КП_2$ – количество правильных ответов в первом и во втором тесте соответственно у данного тестируемого, N – количество тестируемых (сравниваемых пар результатов тестов 1 и 2), а σ_1 и σ_2 – среднеквадратичные отклонения первого и второго теста соответственно. Если валидность по ранговой корреляции можно сразу вычислять на микрокалькуляторе, то валидность по линейной корреляции Пирсона лучше

вычислять с помощью таблицы, куда сначала вносятся разности КП – М для обоих тестов, а затем производится накопление в памяти микрокалькулятора сумм произведений указанных разностей КП – М.

Интерпретация результатов ЛДТ заключается в их анализе для того, чтобы сделать выводы: 1) о контингенте тестируемых в зависимости от целей теста и 2) о самом тесте в зависимости от того, анализируются ли результаты предтеста или окончательного теста.

5.5. СОСТАВЛЕНИЕ ОКОНЧАТЕЛЬНОГО ТЕСТА (ПЕРЕПЛАНИРОВАНИЕ ПРЕДТЕСТА)

На этапе составления окончательного теста, что означает перепланирование предтеста, следует решить, сколько и какие задания можно включить в окончательный тест или во второй предтест. Если заданий с требуемым качеством не хватает, то нужно решить, следует ли использовать несколько одинаковых заданий в разных формах теста или переделать (исправить) некачественные задания. Если меняется хотя бы один отвлекающий ответ, появляется неопределенность в трудности соответствующего задания.

Для более удобного отбора заданий в окончательный тест можно наглядно представить индексы трудности (ИТ) всех заданий путем построения графика роста или убывания трудности по номерам заданий, т.е. задание с самым низким ИТ получает новый номер 1 (отмечают и его действительный номер), за ним следует задание со следующим по величине ИТ (№ 2) и т.д. В идеале желательно иметь прямую линию роста ИТ. Из этого графика видно, сколько и какие задания приемлемы по ИТ.

Трудность и дифференцирующая способность заданий взаимосвязаны по криволинейной зависимости, которая экспериментально определена для лексических заданий теста ИТА-82 и приведена на рис. 9. Контингент испытуемых составлял 400–500 человек для каждой из шести форм теста (390 заданий). Из данных рис. 9 следует, что на пределе приемлемости микроконтекстных заданий (R = 30 %) по трудности ДС имеет весьма большую неопределенность: при R = 30 % ДС может быть от 0,2 до 0,45, т.е. может выйти за пределы приемлемых величин. ДС заданий клоуз-тестов и тестов по грамматике имеет еще больший разброс. Из приведенных данных следует, что трудность 30–35 % является нижним пределом для отбора в тест микроконтекстных заданий.

Трудным представляется вопрос об оценке веса заданий разных частей комплексного теста. С одной стороны, задание по чтению или аудированию заметно сложнее, чем задание по грамматике или лексике, и правильное выполнение первых следовало бы оценивать в 2–3 раза выше. С другой стороны, объективных критериев для определения оценки веса заданий нет. Основой для этого могли бы служить данные факторного и корреляционного анализа, что однако трудоемко. Здесь, видимо, следует прибегать к экспертным оценкам сложности и трудно-

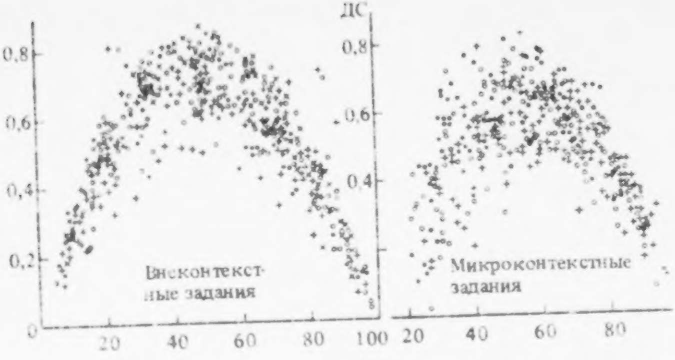

Рис. 9. Зависимость дифференцирующей способности (ДС) лексических заданий теста ИТА-82 от их трудности (% правильных ответов).

сти заданий. Ведь задания по лексике тоже различаются более чем в два раза (от 35 % до 85 %).

6. СОСТАВЛЕНИЕ ТЕСТОВ ПО ЛЕКСИКЕ И ГРАММАТИКЕ

6.1. ОБЩИЕ ПОЛОЖЕНИЯ

Лексические знания и навыки наряду с грамматическими являются субкомпонентами всех компонентов коммуникативной компетенции. Так, модифицируя соответствующие определения, можно сказать, что лексические и грамматические (ЛГ) субкомпоненты лингвистической компетенции (СК № 1) – это навыки и умения образовывать и понимать правильно слова, словосочетания и предложения в соответствии с грамматическими правилами языка. Эти слова, словосочетания и предложения входят в определенные речевые акты, и соответствующая способность их формирования использует ЛГ-субкомпоненты (СК № 2) для образования и понимания лексического наполнения и грамматической структуры указанного речевого акта.

Для образования речевых актов требуются определенные лексико-семантические, синтаксические, контекстуальные и другие когезивные средства для обеспечения связности речевого потока. Умения оперировать указанными средствами и составляют ЛГ-субкомпоненты дискурсивной компетенции (СК № 3).

В речевые акты должны войти такие слова, словосочетания и предложения, которые являются лексически и грамматически адекватными соответствиями в двух языках и культурах в виде ЛГ-субкомпонентов (СК № 4) страноведческой компетенции. Лексико-грамматические суб-

компоненты социолингвистической компетенции (СК № 5) заключаются в способности выбрать такие слова, словосочетания и предложения и с такой грамматической структурой, которые бы подходили к данной социолингвистической ситуации общения. Лексико-грамматические субкомпоненты стратегической компетенции (СК № 6) – это умение находить среди имеющихся (наличных у коммуниканта) слов, словосочетаний и предложений такие, которые обеспечивают оптимальное или минимально достаточное выполнение коммуникативной задачи с учетом ситуации общения и развитости компонентов КК.

Согласно данным таблицы 16 ЛГ-субкомпоненты лингвистической и страноведческой компетенции даже целесообразнее контролировать косвенно из-за экономичности микроконтекстных заданий по сравнению с макроконтекстными заданиями тестов устной речи. С другой стороны, ЛГ-субкомпоненты иллокутивной и стратегической компетенции сложно контролировать косвенно – для них требуется макроконтекст.

Таблица 16

Качественная оценка целесообразности определения лексических и грамматических субкомпонентов коммуникативной компетенции

Принадлежность ЛГ-субкомпонента к компетенции (К)	Характер метода определения		
	косвенный (вне контекста)	полукосвенный (микроконтекстные задания)	непосредственный (макроконтекстные задания)
1. Лингвистической к.	+ +	+	+
2. Иллокутивной к.	–	+	+ +
3. Дискурсивной к.	+	+ +	+ +
4. Страноведческой к.	+ +	+ +	+
5. Социолингвистической к.	+	+ +	+ +
6. Стратегической к.	–	+	+ +

Примечание: «+ +» – наиболее целесообразно, «+» – не очень целесообразно, «–» – нецелесообразно.

Полукосвенный характер при определении ЛГ-субкомпонентов КК имеют тесты чтения, аудирования, грамматические тесты, тесты типа клоуз, диктанты и т.п.

Отметим еще, что при выборе объектов тестирования, приемов и типов заданий следует непременно учитывать используемую методику обучения, программные требования, требования надежности, валидности и экономичности тестов.

6.2. СОСТАВЛЕНИЕ ТЕСТОВ ПО ЛЕКСИКЕ

Методы и приемы для тестирования лексики в рамках СК № 1 лингвистической компетенции сравнительно хорошо разработаны, и эти методы и приемы применимы для тестирования лексики в рамках СК № 4 страноведческой и особенно лингвострановедческой компетенции. Остальные лексические СК тесно взаимосвязаны в контексте устного или письменного общения, и их поэтому лучше контролировать через разные речевые действия в тестах с полукосвенным и непосредственным характером измерения.

Ниже рассматриваются проблемы составления тестов пассивной лексики для чтения и аудирования, поскольку активная лексика входит в указанные выше субкомпоненты №№ 2, 3, 5, 6. Оперирование лексикой для чтения и аудирования предполагает ее быстрое узнавание в ходе этих видов речевой деятельности, которое зависит от синтагматических, парадигматических и денотатных признаков данного слова. Указанные признаки можно для упрощения рассмотрения свести к признакам, связанным со звукографической формой, и семантическим значениям данного слова. По данным Л.В. Банкевича [4], на смешение звукографического образа слова приходится 31,6 % из всех ошибок внеконтекстной семантизации лексики, т.е. семантическое значение имеет большое значение в этом случае.

Рассмотрим подробнее факторы, которые влияют на успешность семантизации слов, т.е. на результат теста. Эти факторы следует учитывать при отборе лексики для теста, чтобы обеспечить репрезентативность материала, а также при составлении заданий лексических тестов (см. рис. 10).

К факторам, связанным со звукографической формой слова, относятся близость, совпадение или несовпадение звуковой и/или графической формы слов ИЯ с этими формами того же ИЯ, т.е. появление возможности внутриязыковой интерференции, или с этими формами родного языка, т.е. появление возможности межъязыковой интерференции. По данным Л.В. Банкевича [4], с последней связано всего лишь 10 % всех ошибок при внеконтекстной семантизации лексики. К факторам, связанным со звукографической формой, относятся еще длина слов и связанная с ней частотность слов как в языке, так и в субъективном опыте учащегося.

К факторам, связанным с семантическим значением слова, относятся такие факторы, как положение данного слова в лексико-семантическом поле и его тематической группе, ассоциативная сила слова и ее направленность (векторный характер), грамматическая форма (часть речи), известен, знаком или неизвестен корень слова, многозначность, словообразовательная ценность, семантическая ценность, сочетаемость с другими словами, а также эмоциональная и стилистическая окрашенность, тематическая значимость (покрываемость разных подъязыков) и другие факторы.

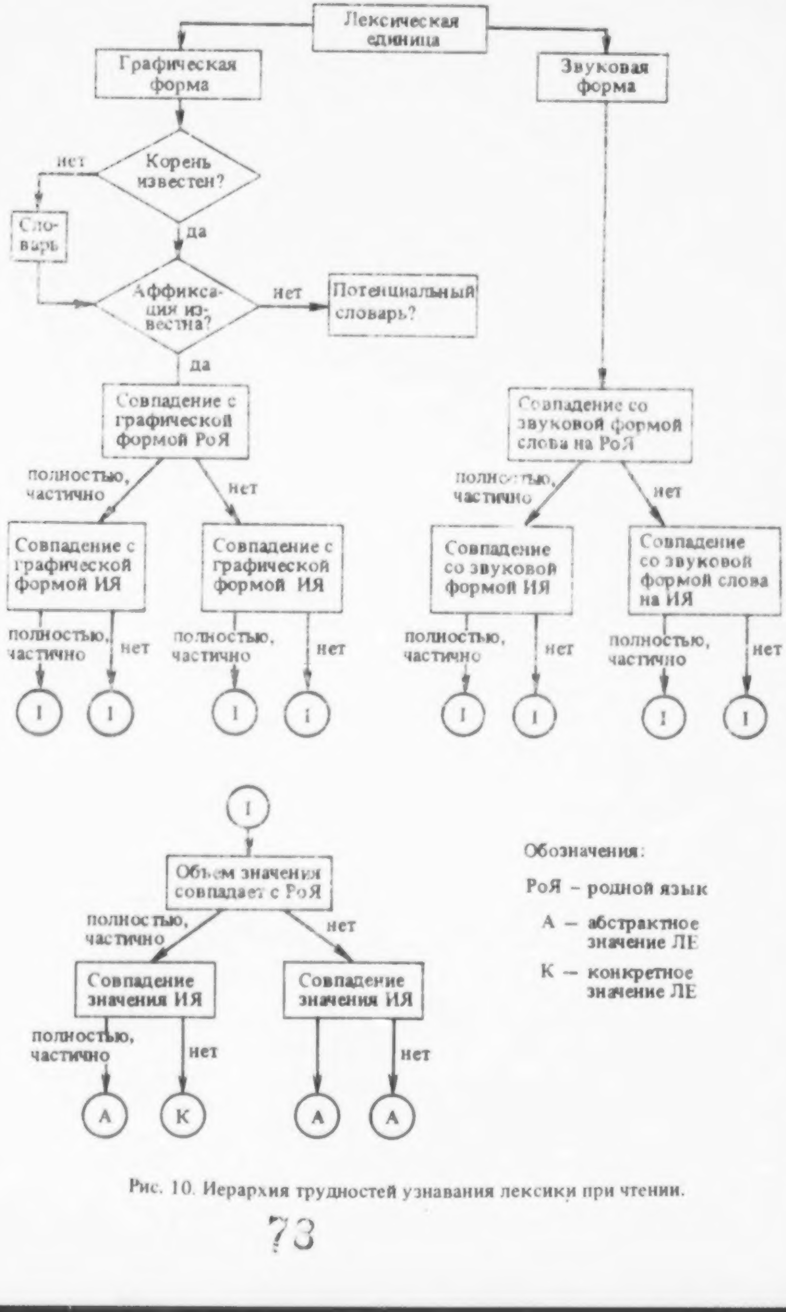

Рис. 10. Иерархия трудностей узнавания лексики при чтении.

Приведенный перечень факторов показывает сложность отбора лексики для теста, который особенно актуален в случае тестов успешности и общего владения ИЯ, а также в случае стандартизованных тестов. И.А. Рапопорт и др. [62] предлагают осуществлять отбор лексики по 7 критериям: 1) односложное – двусложное (или больше слогов); 2) аморфное слово – слово со словообразовательными признаками; 3) одно- или многозначное слово; 4) полное (частичное) совпадение звуковой формы с другими словами на данном ИЯ – несовпадение звуковой формы; 5) полное (частичное) совпадение графической формы с другими словами на данном ИЯ – несовпадение графической формы слова; 6) полное (частичное) совпадение звуковой формы слова ИЯ с РоЯ при разных значениях – полное (частичное) совпадение звуковой формы ИЯ с РоЯ при одинаковых значениях; 7) невысокая частотность слова – его высокая частотность. На основе этих критериев была определена трудность от 0 до 7 баллов для лексических единиц при их отборе для теста.

Во втором случае И.А. Рапопорт и др. [62] использовали методику отбора лексики на основе ее ранжировки по 5 критериям – сочетаемость, семантическая ценность, словообразовательная ценность, многозначность и частотность. Всем этим критериям можно найти количественное выражение в двуязычных словарях, а частотности – в частотных словарях. Чтобы получить один критерий для отбора, для каждого слова по его 5 рангам определяется медианный ранг, который и указывает место данного слова в рассматриваемом списке ЛЕ, предназначенных для отбора.

Можно также использовать методику определения уровня трудности слов, предложенную для немецкого языка М.А. Педановой [53] и для английского языка – И.В. Фельснер [75]. В этой методике уровень трудности слова складывается из баллов, которые суммируются по схеме: 1) имя – 0, глагол – 1; 2) корень известный – 0, неизвестный – 1; 3) конкретное понятие – 0, абстрактное – 1; 4) объем значения на ИЯ совпадает с объемом значения на РоЯ – 0, не совпадает – 1; 5) короткое (1–2 слога) слово – 0, длинное слово – 2.

В ходе экспериментальных исследований в ТПИ установлено, что медианный ранг и составляющие его 5 критериев, кроме частотности, не имеют статистически значимой корреляции с трудностью лексических единиц в тесте. Значимая корреляция трудности ЛЕ обнаружена с повторяемостью ЛЕ в учебниках средней школы (П), частотностью (Ф) и уровнем трудности ЛЕ (У), найденным по методике М.А. Педановой (см. данные табл. 17). Для прогноза трудности R ЛЕ в тесте можно использовать регрессионную формулу в виде:

$$R = A + a \times П + b \times Ф + c \times У, \quad (19)$$

где A, a, b, c являются постоянными, зависящими от характера презентации ЛЕ, параметров контингента учащихся и т.п. Динамический компонент П и трудность для заучивания У, возможно, требуют уточнения и дополнения.

Для обеспечения оптимальной трудности всего теста для данного контингента в тест следует включить как легкие, так и средние по трудности и трудные ЛЕ.

Таблица 17

Корреляция трудности ЛЕ с частотностью, повторяемостью и объективной трудностью усвоения

Название теста	Характер заданий	Количество		Коэффициенты корреляции		
		ЛЕ	студентов	R – П	R – Ф	R – У
№ 3 ТПИ	вне конт. ТМВ	65 ср. шк. и вуз	80	0,49	0,42	0,46
№ 4 ТПИ	микроконт. ТМВ	63 ср. шк.	71	0,40	0,28	0,45
МИРЭА-МЭИ РуЯ-АЯ	вне конт. ТМВ	88 ср. шк.	80	–	0,47	–
ИТА-82 – I	вне конт. СКО	390 ср. шк.	400–500	0,43–0,57	0,35–0,41	–
ИТА-82 – II	вне конт. СКО	390 ср. шк.	400–500	0,47–0,58	0,34–0,39	–
АТУ-85	вне конт. СКО	800 вуз	250–300	0,28–0,31	0,40–0,60	–
АТУ-86	микроконтекстный ТМВ	432 вуз	250–300	0,30–0,36	0,31–0,35	–

Примечание: *Лексика в тестах АТУ составляла около 75 % объема лексического минимума В.В. Морозенко и И.Ф. Турук и корреляция R – Ф получена как по этому словарю, так и по словарю Д. Керрола и Д. Дейвиса.*

Из данных таблицы 17 следует, что все три параметра – П, Ф, У могут быть использованы для отбора лексики для теста. Для теста ИТА-82 лексика отбиралась по ее повторяемости в двух сериях учебников для VIII–X классов средней школы, а в тест АТУ-85 – по повторяемости в словарях учебников и пособий для неязыковых вузов и по

80

частотности в словарях В.В. Морозенко и И.Ф. Турук, а также Дж. Керрола и Д. Дейвиса.

При отборе ЛЕ для теста возникает проблема частей речи. В языке знаменательные слова занимают только около половины всех слов, но включать в тест числительные, предлоги, артикли и т.п. не всегда целесообразно. В таблице 18 приведены сопоставительные данные о трудности разных частей речи при различной технике тестирования. Из этих данных следует, что ни степень контекстуальности заданий, ни тип ответа не оказывают большого влияния на пропорции между трудностями отдельных частей речи, т.е. разная техника тестирования влияет на них примерно одинаково.

Таблица 18

Трудность заданий с разными частями речи при разной технике тестирования

Характер задания и ответа (в скобках приведено общее количество заданий)	Название теста	Части речи (% правильных ответов)			
		имя существительное	глагол	прилагательное	наречие
I. Вне контекста					
1. СКО (390)	ИТА	57,3	50,0	54,3	58,3
2. СКО (24)	ИТА – I ур.	60,5	50,2	53,2	57,0
	ИТА – II ур.	51,0	43,3	36,2	26,5
3. СКО (36)	ТПИ	60,4	68,2	61,5	47,0
4. ТМВ (24)	ТПИ	67,3	72,0	76,0	67,4
II. Микроконтекст					
5. СКО (24)	ТПИ	22,2	15,5	20,7	23,7
6. ТМВ (24)	ТПИ	46,7	36,5	43,0	55,9
III. Макроконтекст (клоуз-тест)					
7. СКО (36)	ТПИ	30,4	34,2	30,0	29,1
8. ТМВ (24)	ИТА – I ур.	64,4	54,8	51,0	69,7
	ИТА – II ур.	48,7	43,0	38,5	45,8
9. СКО (750)	ТПИ	45,5	47,2	35,2	40,1

П р и м е ч а н и е. В таблице 18 сопоставимы (одинаковые ЛЕ и контингенты) ряды 2 и 8, 3 и 7, а также 4, 5 и 6.

При выборе типа заданий обычно предполагают, что характер задания и ответа в нем не влияют на результат теста. Усредненные результаты многих тестов в таблице 19 (ряды 1 и 2) показывают, что ТМВ облегчает

на 25–50 % выполнение заданий теста в зависимости от степени контекстуальности задания, причем задания вне контекста заметно легче, чем задания в микро- и макроконтексте. Микроконтекстом называется здесь одно (иногда 2 коротких) предложение. В тестах ИТА и АТУ близкие трудности заданий в таблице 19 получаются из-за разных типов ответа: вне контекста – СКО, в микро(макро)контексте – ТМВ. В рядах 1 и 2 все тесты имеют одинаковый тип ответа. Задания на завершение или заполнение пропуска по ТМВ заметно легче, чем задания на нахождение синонима или парафразы [48].

Таблица 19

Влияние степени контекстуальности стимула и формы ответа на средний результат (% правильных ответов) лексических тестов

Название эксперимента	Тип ответа	Количество тестируемых	Количество заданий	Контекстуальность стимула			
				вне контекста РоЯ – АЯ	вне контекста АЯ – РоЯ	микроконтекст (завершение)	макроконтекст (клоуз)
1. Клоуз-тест ТПИ	СКО	8×50–60	8×12–24	59,2 %	60,4 %	25,7 %	31,2 %
2. Клоуз-тест ТПИ	ТМВ	8×50–60	8×12–24	–	75,3 %	57,1 %	54,8 %
3. ИТА – I ур.	СКО и ТМВ	400–500	24	–	58,0 %	–	58,7 %
4. ИТА – II ур.	СКО и ТМВ	400–500	24 7	–	38,9 %	–	41,5 %
5. МИРЭА–МЭИ [119]	ТМВ	246 ТПИ	88	74,3 %	75,8 %	–	–
6. АТУ-86	СКО и ТМВ	250–300	300	–	40,2 %	44,3 %	–

Примечание: В тестах 1 и 2, а также 3 и 4 ЛЕ одинаковые.

Данные таблиц 17–19 показывают соотношение некоторых сопоставимых тестовых–нетестовых форм контроля усвоения лексики. Задания

внеконтекстной семантизации со СКО являются письменными переводами слов. Различия между данными рядов 1 и 2 в таблице 19 относятся во многом к подсказке, которую дает ТМВ. При этих сравнениях были использованы только те задания, у которых ДС > 0,40, чтобы исключить влияние содержания текста задания, неудачных отвлекающих и т.п.

Вышеуказанное заметно усложняет выбор типа заданий, поскольку почти все из 12 основных типов заданий подходят для тестирования лексики, но они отличаются, кроме прочего, еще своим влиянием на результат теста. Основным критерием для выбора типа заданий лексических тестов является соответствие задачам и методике обучения, задачам теста и требованиям обеспечения надежности, валидности и экономичности теста.

При составлении лексических заданий следует решить, на основе какой лексики будут составлены тексты стимулов (основы) заданий. Чтобы исключить возможность появления в задании еще одной проблемы («пропуска») в виде незнакомого или недостаточно усвоенного слова, надо использовать более легкую лексику или проверять соответствующие знаменательные слова текста задания. Например, в тестах ИТА I уровня (для IX–X классов) лексика была выбрана из списка ЛЕ VII–VIII классов.

Приведем примеры основных типов заданий по лексике.

1. Задания вне контекста по ТМВ

1.1. Сопоставление ЛЕ

Пример из АТУЛЕКС-87.

Part I. Synonyms

C-7 to receive – 1) to leave, 2) to believe, 3) to belong; 4) to get

Part III. Antonyms

C-33 present – 1) absent; 2) abstract; 3) except; 4) expect

Могут быть сопоставлены также списки слов на ИЯ, предложения или их части. Из-за ограниченного объема пособия невозможно описать лексические задания с сопоставлением, в которых используются картинки и другие виды зрительной наглядности.

1.2. Задание в виде группировки – исключите слово, означающее самое меньшее:

1) inch, 2) yard, 3) mile, 4) foot

2. Задания с микроконтекстом по ТМВ

2.1. Завершение

Для лексических заданий тестов ИТА и АТУ обнаружена слабая, но статистически значимая корреляция трудности заданий с длиной задания в словах и с местом пропуска в предложении: чем дальше пропуск, тем труднее задание. Однако не рекомендуется использовать в качестве места пропуска первое слово в задании.

Пример из ИТА-83 форма С-2.
С-46
In autumn some birds ... to the south.
1) fly 2) jump 3) walk 4) run

2.2. Задание на определение синонима
IEA UNESCO N 26
They usually have **fruit** here.
A. Kind of food; B. Music; C. Lessons

2.3. Интерпретация
I shall be back soon. It is ... .
1) promise; 2) recommendation, 3) advice

Подробное обсуждение техники тестирования лексики приведено в пособии Л.В. Банкевича [4].

6.3. СОСТАВЛЕНИЕ ТЕСТОВ ПО ГРАММАТИКЕ

При дискретном подходе к составлению тестов по лексике и грамматике теоретической основой служили положения структурной и трансформационной грамматики – владение лексическими значениями и грамматическими формами (ГФ) проверялось на уровне морфем, слов, словосочетаний и предложений. При интегративном подходе к этим уровням прибавляется еще уровень связного текста, дискурса и сверхфразового единства (СФЕ). Чем выше коммуникативная компетенция (КК) учащихся, тем больше проверка на указанных первых уровнях вливается в уровень текста, дискурса, т.е. владение ГФ проверяется в макроконтексте.

Если в начале развития коммуникативно-ориентированной методики обучения ИЯ методисты пренебрегали грамматической точностью и правильностью (лишь бы сообщение было понято), то с 80-х годов больше внимания обращается и на грамматическую точность устной речи. Однако в чистом виде в макроконтексте невозможно выделить лексические и грамматические явления из-за единства и сложности семантических и синтаксических связей в структуре текста. Поэтому на начальном и среднем этапах обучения традиционные способы проверки владения лексикой и грамматикой сохраняют свое значение.

Кроме лингвистической компетенции грамматические навыки имеют большое значение в таких субкомпонентах КК, как СК № 3 дискурсивной компетенции (синтаксические когезивные средства) и СК № 5 социолингвистической компетенции (сокращенные, нерегулярные ГФ, пропуски и т.п.). Последние лучше контролировать полукосвенными приемами в макроконтексте, и ниже будут обсуждены проблемы составления тестов пассивной грамматики для чтения и аудирования. Формы активной грамматики составляют содержание СК № 3 и № 5, которые вкратце обсуждались выше.

Обычно считается, что отбор объектов тестирования по грамматике легче, чем по лексике. Однако и здесь возникают специфические проблемы. Для отбора ГФ для теста обычно исходят из: 1) пройденного материала в учебнике (ПТ), требований программы и грамматического минимума (тест достижений). Последний, однако, недостаточно разработан с точки зрения трудности отдельных ГФ; 2) частотности ГФ в языке (подъязыках) и их повторяемости в учебнике (в субъективном опыте учащихся), которые недостаточно исследованы. Например, между повторяемостью ГФ в учебниках средней школы и результатами тестов ИТА не обнаружено статистически значимой корреляции [136]; 3) трудности отдельных ГФ по данным контрастивной лингвистики, полученным путем сопоставления изучаемых ГФ данного ИЯ с соответствующими ГФ родного языка учащихся; 4) трудности отдельных ГФ для узнавания при чтении и трудности усвоения, о которой имеются только теоретические соображения. Попытка систематизации последних приведена на рис. 11.

При узнавании ГФ некоторое интерферирующее влияние (дополнительная трудность) может оказывать и уровень I на рис. 11, но возникающая трудность скорее связана с общей абстрактности ГФ. Уровни трудности I–IV относятся к внешней форме рассматриваемой грамматической категории, а функциональное значение (ФЗ) ГФ на ИЯ и степень и объем его совпадения с соответствующим ФЗ на родном языке (уровни V–VIII) относятся к внутренней структуре ГФ.

При составлении изображенной на рис. 11 иерархии предполагалось, что если данная ГФ имеет такое ФЗ, которое полностью совпадает с ФЗ соответствующей ГФ на родном языке, то эту ГФ можно считать конкретной. Кроме того, если ГФ многокомпонентна, то вероятность нескольких ФЗ на ИЯ весьма мала. Для упрощения анализа роль контекстуальных связей, порядка слов и т.п. опущена.

При выборе типа заданий и составлений заданий тестов по грамматике следует исходить из следующих принципов: 1) задания должны отражать не столько знания правил образования, сколько оперирование с ГФ (употребление, узнавание и т.п.); 2) каждый объект тестирования (ГФ) имеет свои наиболее подходящие типы заданий. Например, предлоги, местоимения, артикль и т.п. лучше тестировать в виде сопоставления нескольких (3–5) стимулов (минитекста) со списком вариантов ответов; 3) поскольку в тестах по грамматике ответы обычно достаточно однозначны, то для преподавательских тестов более рациональны задания со СКО или СКО полуоткрытого типа, позволяющие лучше выявить типичные ошибки; 4) текст задания должен содержать только хорошо усвоенную лексику, обеспечивающую при оптимальной средней длине заданий 9± 1–2 слова достаточную ситуативность и однозначность стимула (основы) задания; 5) использование перевода в заданиях оправдано только на начальном этапе обучения и в трудных по данным контрастивной лингвистики случаях; 6) на одну ГФ желательно иметь 2–3 задания в одной форме теста.

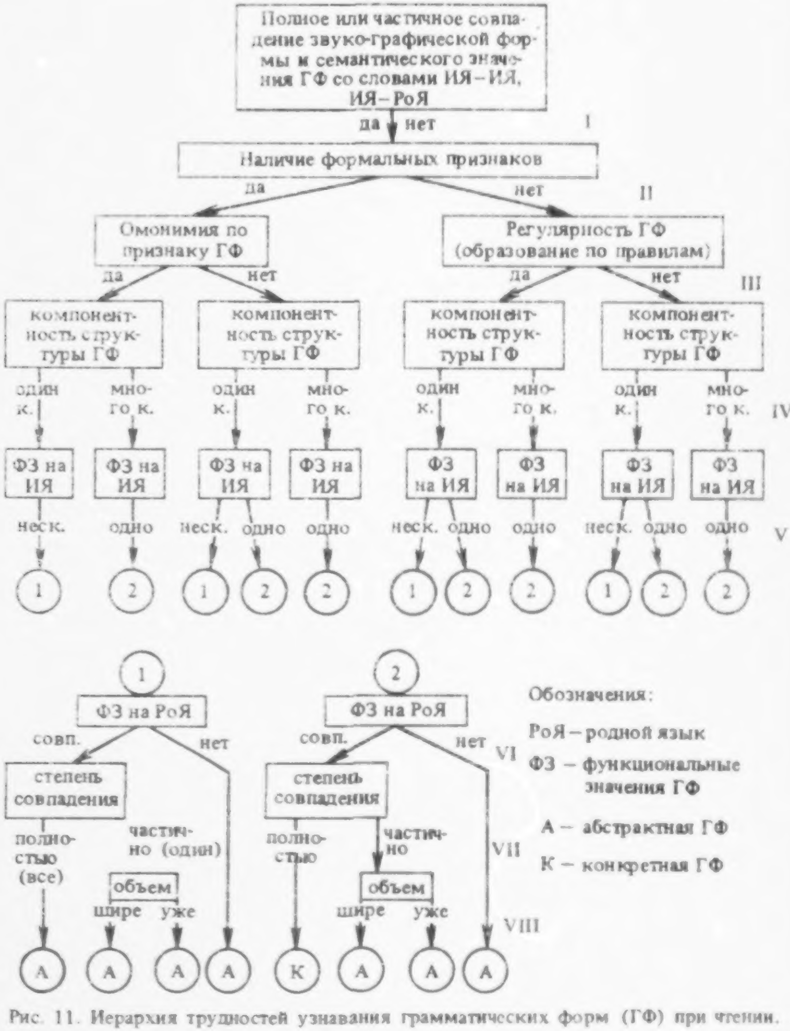

Рис. 11. Иерархия трудностей узнавания грамматических форм (ГФ) при чтении.

Ниже приведено несколько примеров типов заданий по грамматике.
Микроконтекстные задания по ТМВ

1. Завершение, ИТА-84 С-1 № 18:
He can ... English and German.
1) to speak; 2) speak; 3) speaks; 4) speaking

2. Трансформация по ТМВ или со СКО (лучше):

This way is not more difficult. It is ... .

1) easiest; 2) easier; 3) easy; 4) easter

3. Порядок слов: указать правильное место наречия в предложении (по ТМВ или со СКО):

The work (1) should (2) be (3) done (4) well. (always)

7. СОСТАВЛЕНИЕ ТЕСТОВ ПО ЧТЕНИЮ И АУДИРОВАНИЮ

7.1. ОБЩИЕ ПОЛОЖЕНИЯ

Тесты по чтению и аудированию относятся к традиционным речевым тестам, с помощью которых проводится оценка таких показателей понимания при чтении или аудировании, как полнота, глубина, точность, отчетливость и другие. Указанная оценка имеет косвенный характер, так как степень понимания можно установить только посредством какой-то другой деятельности, например, пересказа, ответов на вопросы и т.п. Степень опосредованности наименьшая, если после чтения следует нарисовать схему, показать путь по карте и т.п., и наибольшая, если за чтением последует, например, пересказ на том же иностранном языке.

При чтении и аудировании имеет место такая коммуникация, при которой обратная связь ограничена. Читающие (слушающие) обычно не могут при необходимости запрашивать дополнительную информацию, чтобы выяснить окончательно и достоверно коммуникативное намерение автора; читающий (слушающий) должен полагаться на свою стратегическую компетенцию при раскрытии смысла сообщения, используя при этом механизмы вероятностного прогнозирования (контекстуальной догадки).

По сравнению с другими компонентами КК роль стратегической компетенции при чтении и аудировании больше всего отличается от таковой при говорении. Она заключается в способности выявления коммуникативных намерений автора сообщения путем построения гипотез о достоверности того или иного коммуникативного намерения автора. В тесте роль этого компонента КК выявляется с помощью вопросов об основной мысли текста, об общем содержании, о ходе доказательства и т.п.

Дискурсивная компетенция (ДК) при чтении и аудировании проявляется в способности узнавания и использования для смысловой переработки информации текста различных когезивных средств, в том числе между предложениями и между абзацами. Для выявления ДК можно использовать вопросы на определение понимания смысловых связей между денотатами в предметно-тематической структуре текста и между уровнями предикации в логико-смысловой структуре текста.

Роль иллокутивной компетенции (ИК) при чтении и аудировании заключается в смысловой обработке информации в речевых актах в соответствии с их иллокутивной силой, чтобы раскрыть предметное содержание текста. В тексте роль этого компонента выявляется с помощью вопросов к смысловым вехам (кускам), к ключевым предложениям в сверхфразовых единствах (СФЕ), а также по пониманию основных фактов и важных деталей текста.

Социолингвистическая компетенция при чтении и аудировании проявляется в виде способности узнавания и использования социолингвистических и стилистических вариаций лексико-грамматических и фонологических форм текста для смысловой переработки содержащейся в нем информации. Страноведческая компетенция при чтении и аудировании проявляется в виде способности узнавания и интерпретации страноведческой и лингвострановедческой информации текста. Последние компоненты КК выявляются в основном на уровне предложения и абзаца (СФЕ).

Для определения развитости коммуникативной компетенции можно расположить данные по ее компонентам согласно определенной схеме или построить диаграмму всех или нескольких ее компонентов, которая показывает развитость этих компонентов КК по единой критериальной шкале – владеет или не владеет читающий/ слушающий теми или иными навыками и умениями чтения или аудирования, которые связаны с соответствующими компонентами КК. В приведенной методике составления речевых тестов имеется еще ряд нерешенных проблем, и ее можно использовать только в том случае, если обучение ИЯ ведется соответственно ей.

Все факторы, влияющие на понимание при чтении и аудировании, зависят: 1) от характеристик текста и модальности его предъявления (слуховая или зрительная); 2) от взаимодействия текста с читающим/ слушающим; 3) от индивидуально-психологических особенностей читающего/слушающего (возраст, степень владения ИЯ и т.п.).

Отбор текстов для тестов по чтению и аудированию в зависимости от их характеристик освещен в литературе весьма поверхностно. Проблем, однако, довольно много, так как текст следует рассматривать с точки зрения множества факторов и характеристик, которые приведены на рис. 12 и которые позволяют достоверно определить методическую ценность и трудность текстов для их сравнения и отбора для тестов. Рассмотрение проблем отбора текста следует ниже (отдельно для тестов по чтению и аудированию).

Составление заданий для тестов по чтению и аудированию сложнее, чем в случае тестов по лексике и грамматике, так как эти задания требуют большего объема и более сложной мыслительной деятельности тестируемых, а с другой стороны, имеется больше возможностей угадать правильный ответ (при СКО тоже) по формальным признакам стимула, вариантов ответов, из контекста и т.п.

Общее требование к заданиям тестов по чтению и аудированию состоит в том, что они должны соответствовать объектам тестирования, целям

ТЕКСТ

I. Факторы содержания текста

1. Тематика
   - 1.1 релевантность
   - 1.2 познават. ценность (новизна, знакомость, интересность)
2. Тип речи
   - 2.1 описательный
   - 2.2 повествовательный
   - 2.3 объяснительный (рассуждение, доказательство и т.п.)
3. Жанр
   - 3.1 релевантность
   - 3.2 относит. трудность
   - 3.3 степень коммуникативности
4. Функциональные стили
   - 4.1 релевантность
   - 4.2 стилистич. сложность
   - 4.3 сложность контекста и подтекста
5. Сложность
   - 5.1 Абстрактность понятий, фактов и т.д.
   - 5.2 Информационная насыщенность

Примечание 1: Релевантность означает соответствие целям общения, программы, теста, а также возрасту и потребностям обучаемых

II. Факторы структуры текста

1. предметно-тематическая структура
2. логико-смысловая стр-ра
3. логико-композиционная стр-ра
4. развитость ориентировочной основы

III. Лексико-грамматические факторы

1. трудность ЛЕ
2. трудность ГФ
3. частотность знаменательных слов
4. длина предложений и слов (читабельность)
5. Количество незнакомых ЛЕ и ГФ
6. Соотношение знаменательных и функциональных слов
7. Аутентичность языка

Примечание 2: Структурно-композиционные факторы п. 1–4 характеризуются:
- релевантностью количеством структурных элементов (фактов)
- их распределением и плотностью в тексте
- количеством и развитостью уровней в предикативной структуре

IV. Факторы коммуникативных характеристик текста

1. Сложность речевых (коммуникат.) актов
2. Степень и количество требуемых компонентов КК
3. Насыщенность социолингвистич. вариациями и их характер
4. Характер страноведческой информации (знакомость, релевантность и т.п.)
5. Дискурсивная структура
   - 5.1 Маркированность и характер когезивных средств
   - 5.2 Характер паралингвистических маркеров
6. Сложность и разнообразие требуемых стратерий чтения (контекстуальная догадка)

V. Факторы модальности

1. При чтении - видимость (качество бумаги и печати) и удобочитаемость (шрифт)
2. При аудировании
   - темп речи
   - качество звука
   - тембр, дикция и т.п.

Рис. 12. Факторы текста, влияющие на понимание при чтении и аудировании.

и задачам теста, виду чтения или аудирования и методике обучения чтению и аудированию. Кроме того, важно стремиться к возможно большей коммуникативности заданий, отдавая предпочтение проверке понимания (интерпретации, а не узнаванию) более обобщенных объектов (идей, тем, основных мыслей, главных фактов), причем в основном на уровне всего текста, его частей, абзацев, а не предложений, слов и словосочетаний.

Стимулы заданий должны отвечать следующим основным требованиям: 1) в них следует использовать простые формулировки, простые лексико-грамматические формы, чтобы облегчить понимание стимула, в некоторых случаях стимул и ответ лучше представить на родном языке тестируемых; 2) стимул желательно составлять не словами текста, а в виде парафразы и реконструкции слов и предложений текста; 3) независимость от других заданий (их стимулов и отвлекающих) и зависимость от данного текста, чтобы было невозможно ответить правильно без прочтения текста теста. При этом следует учесть возможные предварительные знания тестируемых о предмете текста данного теста.

В случае использования ТМВ отвлекающие можно найти среди неправильных ответов предтеста, в котором тестируемым предъявлялись задания со СКО. Отвлекающие ответы могут быть внеконтекстными или контекстными, и для их составления применимы те же принципы, которые используются в случае лексических тестов (см. с. 64). Отвлекающие должны отвечать следующим основным требованиям: 1) в них нельзя использовать те же слова, которые использованы на соответствующем месте текста; 2) нельзя использовать странные, явно бессмысленные отвлекающие; 3) отвлекающие должны иметь примерно одинаковую длину в словах или слогах и примерно одинаковую степень обобщенности; 4) отвлекающими могут быть слова или словосочетания, которые по своему местоположению лежат близко к месту объекта стимула, но которые по своему смыслу нерелевантны; 5) отвлекающими могут быть слова, словосочетания или предложения, отражающие либо неправильное, неточное толкование или понимание идей, тем, фактов текста либо несоответствие, модификации идей, тем, фактов. При этом должна быть обеспечена однозначность (единственность) правильного ответа, т.е. соблюдено достаточное контекстуально-смысловое расстояние между правильным ответом и отвлекающим.

7.2. СОСТАВЛЕНИЕ ТЕСТОВ ПО ЧТЕНИЮ

Умения и навыки чтения подразделяются на две взаимосвязанные группы — умения перцептивной переработки воспринимаемого языкового материала и умения смысловой обработки извлекаемого в ходе чтения информации. Объектами тестирования для первой группы служат такие навыки и умения, как степень усвоения алфавита, сформированности звуко-буквенных ассоциаций и другие, которые не имеют большого значения на вузовском уровне подготовки по ИЯ.

Объектами тестирования для второй группы служат такие умения, как, например, понимание общего смысла текста, выделение основной мысли, главной идеи текста, понимание всех фактов и т.п. Основные объекты тестирования для разных видов чтения приведены в таблице 20. Отметим, однако, что составление тестов ознакомительного и просмотрового (поискового) чтения затруднительно технически из-за того, что в них следует ограничивать время чтения так, чтобы каждый тестируемый прочел бы текст только один раз с определенным темпом. Поскольку обычно тест воспринимается тестируемыми как весьма серьезное дело, то появляется опасность, что они читают текст для ознакомительного чтения со скоростью изучающего чтения. Кроме того, темпы у разных тестируемых разные, что усложняет групповой метод проведения теста, так как после прочтения текста следует его закрыть и перейти к проверке понимания прочитанного. В крупных стандартизованных тестах явно тестируется в основном изучающее чтение, поскольку на одно задание (вместе с чтением в среднем 150–200 печ. зн. текста) отводится приблизительно 1 минута (см. табл. 10, 11). На ответ требуется 0,5–0,7 минут, и более сильная часть тестируемых вполне успевает два раза прочесть текст для выбора правильного ответа.

Таблица 20

Объекты тестирования при разных видах чтения

Объекты тестирования (умения)	Виды чтения		
	изучающее	ознакомительное	поисковое и просмотровое
1. Понимание общего смысла текста	+ +	+ +	+
2. Понимание главной идеи, основной мысли	+ +	+ +	+
3. Понимание всех фактов, выводы по ним	+ +	–	–
4. Понимание некоторых основных фактов	+ +	+ +	+
5. Ориентация в логико-смысловой структуре	+ +	+	–
6. Понимание деталей и подробностей в тексте и в иллюстрациях	+ +	+	–

Продолжение

Объекты тестирования (умения)	Виды чтения		
	изучающее	ознакомительное	поисковое и просмотровое
7. Установление причинно-следственных и других смысловых связей в тексте (глубина понимания)	+ +	+	–
8. Оценка информации, разделение основной и второстепенной информации	+ +	+ +	–
9. Определение значимости содержащейся в тексте информации	+	+	+ +
10. Выявление нужной информации	+	+	+ +
11. Использование ориентировочной основы	+	+	+ +
12. Интерпретация, выводы, оценка	+ +	+ +	+

Примечание: «+ +» – необходимость; «+» – желательность; «–» – нецелесообразность.

Таблица 21

Оценка возможностей использования разных типов заданий для разных объектов тестирования

Типы задания и ответа	Объекты тестирования по № в табл. 20											
	1	2	3	4	5	6	7	8	9	10	11	12
1. Формулировка идеи, темы, плана и т.п. (СКО)	+	+	+	+	–	–	–	–	+	+	–	+
2. Перевод – частичный	+	+	+	+	+	+	+	–	–	–	–	–
– полный	+	+	+	+	+	+	+	–	–	–	–	–
3. Мыслительные задачи (СКО) – объяснить, нарисовать и т.п.	+	+	+	+	+	+	+	+	+	+	–	+
4. Устный или письменный пересказ (изложение) на РоЯ	+	+	+	+	+	+	+	+	+	+	–	+
5. Вопросы по содержанию по ТМВ	–	–	+	+	–	+	–	–	+	+	–	–
со СКО	+	+	+	+	+	+	+	+	+	+	+	+
6. Серия утверждений (интерпретация)	–	–	+	+	–	+	–	–	+	+	–	–

Типы задания и ответа	Объекты тестирования по № в табл. 20											
	1	2	3	4	5	6	7	8	9	10	11	12
7. Утверждение (интерпретация) с завершением по ТМВ	–	–	+	+	–	+	–	–	+	+	–	–
со СКО	+	+	+	+	+	+	+	+	+	+	+	+
8. Вопрос и утверждение с завершением по ТМВ	–	–	+	+	–	+	–	–	+	+	–	–
со СКО	+	+	+	+	+	+	+	+	+	+	+	+
9. Группировки слов, предложений, абзацев (ТМВ)	+	+	+	+	–	+	–	–	+	+	+	+
10. Задания (действия) с наглядными ответами (рисунок, карта, схема) по ТМВ	–	–	+	+	–	+	–	–	+	+	–	–
со СКО	+	+	+	+	+	+	+	+	+	+	–	+

Примечание: «+» – сочетание возможно и желательно; «–» – либо невозможное сочетание, либо трудно качественно составить задание.

Речевые тесты могут быть составлены с разной степенью коммуникативности в зависимости от того, каковы объекты тестирования (общее понимание, понимание деталей, на уровне смысла или слова и т.п.) и сознательно ли к ним применяется подход КК. Поэтому для увеличения степени коммуникативности речевых тестов следует использовать более обобщенные объекты тестирования и задания с ответами полуоткрытого типа или со СКО.

На практике при составлении тестов чтения используются приведенные в таблице 22 критерии отбора текстов. В текстах объемом 1–2 стр. соотношение знаменательных и служебных слов (З/С) может колебаться по данным 24 текстов до трех раз, что отражается на их методической ценности (при З/С = 1 : 2) и на их трудности (при З/С = 3 : 1). Коэффициент информационной насыщенности (ИН) определяется как отношение числа примененных лексем к общей длине текста или числу его словоупотреблений, показывая этим повторяемость слов в данном тексте.

В случае гетерогенного по уровню подготовки контингента показатель % незнакомых слов (НС) довольно неопределенен, поскольку количество НС для более сильной половины контингента может быть в два-три раза меньше, чем для более слабой половины. Кроме того, часть НС может быть в потенциальном словаре учащихся, а часть может быть

Количественные значения некоторых критериев отбора текстов для тестов чтения

Название критериев	Виды чтения		
	изучающее	ознакомительное	просмотровое, поисковое
1. Тип речи в тексте	все типы	описательный, повествовательный	описательный, повествовательный
2. Незнакомые слова (%)	5–10	3–5	3–5
3. Читабельность по Флешу более	25	35	45
4. Соотношение знаменательных и служебных слов (З/С) максимально (%)	65 : 35	60 : 40	60 : 40
5. Информационная насыщенность (ИН) по повторяемости лексем	0,4	0,3	0,3
6. Объем текста максимально в печ. зн.	500–1500	1500–4500	1500–4500

понята с помощью контекстуальной догадки. Определение НС может быть произведено другой выборкой учащихся путем выписывания НС, тогда средний % НС несколько ниже его вероятного истинного значения, или группой преподавателей, тогда средний % НС несколько выше его вероятного истинного значения. Следовало бы учитывать и количество денотатов и их структуру (распределение) в денотатных графах.

Важным индикатором развитости навыков чтения является темп или скорость чтения. Чтение вслух контролируется увереннее и может быть использовано в тестах для определения произносительных навыков. При увеличении количества текстов, а также длительности чтения до 3–4 мин. между скоростью чтения вслух и про себя нами обнаружена корреляция до 0,90. Зависимость между скоростями чтения вслух (В) и про себя (С) имеет криволинейный характер, поскольку максимальная скорость чтения вслух ограничена. Однако для средне-гетероген-

ных контингентов данная часть кривой аппроксимируется прямой линией по формуле:

$$B = A + KC, \tag{20}$$

где А и К – постоянные, зависящие от характеристик текста и контингента. В случае чтения научно-технической литературы постоянная А меньше и К больше, т.е. чтение про себя идет медленнее. Порядок величины А – 150–250 печ. зн./мин. и К – 0,4–0,6 для родного и английского языка [32].

Таблица 23

Экспериментальные данные о связи между скоростью чтения и трудностью текста у первокурсников ТПИ

Показатели	Чтение вслух			Чтение про себя		
	ЭЯ	РуЯ	АЯ	ЭЯ	РуЯ	АЯ
Скорость чтения (печ. зн./мин.):						
Студенты-русские	–	1150	560	–	1760	650
Студенты-эстонцы	1000	–	660	1360	–	820
Коэффициенты корреляции:						
1. Вслух – про себя	–	–	–	0,70	0,63	0,80
2. Роя – АЯ	0,63	0,40	–	0,62	0,47	–
3. С % незнакомых слов	–	–	0,36	–	–	0,59
4. С коэффициентом читабельности Флеша	–	–	0,45	–	–	0,45
5. С общими результатами теста № 5 ТПИ	–	–	0,55	–	–	0,60
6. С результатами части чтения теста № 5 ТПИ	–	–	0,45	–	–	0,42

В таблице 23 все данные получены при чтении одного текста в течение 1 минуты. Взяты средние значения скорости чтения за 5 лет у студентов, поступивших в ТПИ (от 80 до 500 человек ежегодно). Уровень подготовки студентов-русских ТПИ несколько ниже стандартных средних по тесту ИТА-84 II уровня и студентов-эстонцев несколько выше стан-

дартных средних по тесту ИТА-84 I уровня. Тест №5 в таблице 23 был предтестом стандартизованного в ЭССР теста АТ-78 и включал 89 заданий, из них 8 по чтению.

7.3. СОСТАВЛЕНИЕ ТЕСТОВ ПО АУДИРОВАНИЮ

Тестами по аудированию называются такие тесты, с помощью которых производится оценка просодических, лексических и грамматических навыков и степени понимания при слуховом восприятии речевого сообщения. В случае дискретного подхода развитость указанных навыков определяется с помощью слов, словосочетаний и предложений вне макротекста. В случае интегративного подхода развитость указанных навыков определяется вместе с определением степени понимания при аудировании.

По сравнению с чтением восприятие и понимание сообщений при аудировании труднее из-за таких причин, как меньшая пропускная способность слухового канала, заданный (иногда неприемлемый) темп восприятия, невозможность регрессии, большая нагрузка всех видов памяти слушающего, большая роль вероятностного прогнозирования и др. Из-за этого составление доступного по трудности и качественного теста аудирования сложнее и требует больше опыта и больше предварительного опробования.

Основные факторы, которые влияют на успешность аудирования – на степень понимания (глубину, точность и т.п.), приведены на рис. 13. В таблице 24 приведены рекомендуемые в методической литературе предельные значения некоторых показателей процесса аудирования (см. также табл. 11), которые могут служить ориентирами при составлении тестов по аудированию.

На начальном и среднем этапе обучения в диагностические (преподавательские) тесты по аудированию могут быть включены просодические, лексические и грамматические навыки в равной пропорции. Просодические навыки подразделяются на дифференциацию звуков и фонем, словесное, фразовое и логическое ударение и интонацию в предложении. В основном используются задания по ТМВ или двойной выбор (разный – одинаковый). Используются также задания типа: какая пара слов совпадает – тестируемый слышит одно или несколько слов и сопоставляет их с несколькими или одним словом, напечатанном в стимуле задания. К просодическим навыкам относятся еще расчленение речевого потока на синтагмы в опоре на ритмико-мелодическую организацию предложения, вычленение главного и второстепенного в сообщении в опоре на логическое ударение и паузацию, а также распознавание коммуникативного типа предложения по его ритмико-мелодической организации, для которых используются задания со СКО и по ТМВ.

Просодические навыки образуют отдельный просодический субкомпонент лингвистической компетенции наряду с лексическими и грамматическими субкомпонентами. Определение развитости компонентов КК

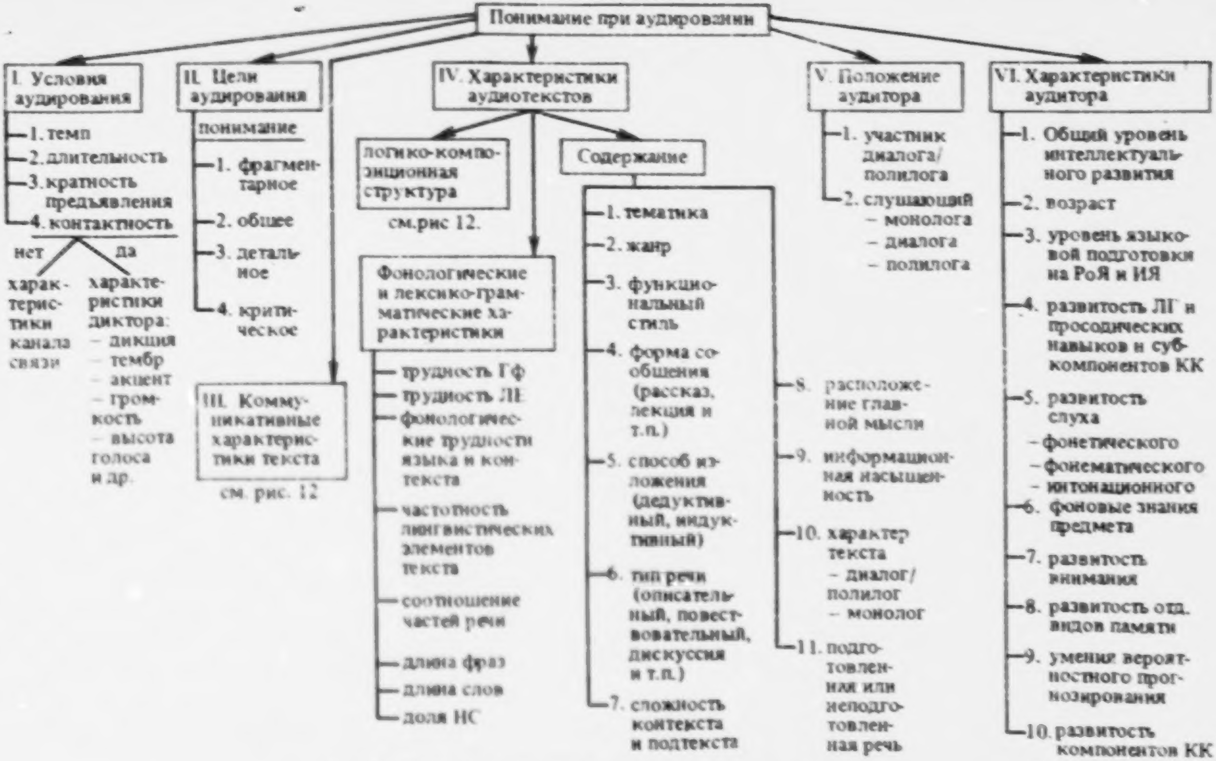

Рис. 13. Факторы, влияющие на понимание при аудировании.

Таблица 24

Количественные значения некоторых показателей процесса аудирования

Показатели	Монологическая речь	Диалогическая речь
Максимальный темп аудирования (слов/мин.)	120–130	125–140
Максимальная длительность звучания (мин.)	2–3	до 2
Максимальная длина текста в тестах (слов)	250–375	250–280
Средняя длина текста в тестах (слов)	110	–
Максимальная длина фраз (слов)	до 11–12	до 10
Оптимальная доля незнакомых слов (%)	2–3	до 2
Максимальное количество смысловых кусков	5	5

при аудировании производится во многом аналогично чтению с учетом специфики процесса аудирования.

При тестировании лексико-грамматических навыков аудирования можно использовать наглядные средства (картинки, карты, схемы, слайды, кинофрагменты и т.п.) и разные действия с ними, а также физические действия (выполнение команд, просьб и т.п.).

В таблице 25 приведены основные объекты тестирования и рекомендуемые к ним типы заданий в тестах аудирования. Из-за сложного характера процесса аудирования задания по ТМВ предложены и для проверки понимания темы, главных мыслей и идей. При выборе типа задания следует в основном исходить из соображений экономичности и степени коммуникативности теста.

Обычно стимул заданий по ТМВ или со СКО напечатан на листах заданий, в этом случае тестируемый может успеть с ним познакомиться до прослушивания текста. Вместо этого можно предъявить стимул устно или письменно после текста, особенно в том случае, если заданий к тексту мало. Следует стремиться к тому, чтобы к каждому аудиотексту было составлено хотя бы две формы (вариантов) заданий, чтобы исключить «сотрудничество» сидящих рядом тестируемых. На начальном и среднем этапе обучения лучше использовать задания на родном языке, а на более продвинутых этапах – на иностранном.

Таблица 25

Объекты тестирования и основные типы заданий в тестах аудирования

Объекты тестирования при аудировании	Формы предъявления стимула			
	устная + письменная	устная с наглядными средствами	устная с действиями	письменная
1. Просодические навыки	3, 4, 5	3, 4, 5	–	3, 4, 5
2. Лексические навыки	1–11	3–6	3, 10	1–11
3. Грамматические навыки	1–11	3–6	3, 10	1–11
Умения понимания речи на слух. 4. Определение темы сообщения	1–3, 6–14	1–3, 6–14	7–11, 12–14	1–3, 6–14
5. Понимание главных мыслей	1–3, 6–14	1–3, 6–14	7–11, 12–14	1–3, 6–14
6. Выделение и группировка фактов	1–14	1–14	7–14	1–14
7. Понимание деталей и второстепенной информации	1–14	1–14	7–14	1–14
8. Выделение иллюстрирующей и аргументирующей информации	1–14	1–14	1–14	1–14
9. Установление смысловых частей (кусков) высказывания	7–12 12–14	–	7–11 12–14	7–11 12–14
10. Установление причинно-следственных и других смысловых связей	7–11 12–14	–	7–11 12–14	7–11 12–14
11. Понимание подтекста и эмоционально-оценочной информации (критическое понимание)	1–14	1–14	–	1–14
12. Антиципация содержания	1–14	1–14	–	1–14

Примечание: Типы заданий по таблице 14: 1 – завершение, 2 – синоним (парафраза), 3 – ДВ, 4 – группировка, 5 – сопоставление, 6 – интерпретация (все по ТМВ), 7 – СКО, 8 – завершение со СКО, 9 – ответ на вопрос, 10 – интерпретация, 11 – парафраза, 12 – исправление (9–12 все со СКО), 13 – мыслительные задачи, 14 – пересказ (устный или письменный).

7.4. КОМБИНИРОВАННЫЕ РЕЧЕВЫЕ ТЕСТЫ

Комбинированными речевыми тестами называются тесты, которые нацелены на проверку умений в нескольких видах речевой деятельности. Жизненно реальными комбинациями являются: 1) чтение – говорение (устная передача содержания); 2) чтение – письмо (письменное изложение содержания на ИЯ); 3) аудирование – говорение (участие в дискуссии, интервью и т.п.).

Особенностью комбинированных тестов является то, что в них используются задания, позволяющие с максимально возможной эффективностью проверять навыки и умения во всех рассматриваемых видах речевой деятельности. Очевидно, что составлять задания по ТМВ очень трудно, и такие задания ниже не рассматриваются.

В данном случае центральным вопросом является максимально возможная объективизация и однообразие (регламентация) процедуры оценки СКО. Можно использовать следующие разновидности шкал оценки результатов теста: 1) нормо-ориентированную (%, баллы и т.п.); 2) критериальную, например, по критерию: выполнение требований программы, т.е. выполняет или не выполняет тестируемый требования программы по рассматриваемым объектам тестирования; 3) комбинированная шкала из перечисленных, например, зачет по какому-то количеству ошибок, если их больше – то незачет и т.п.; 4) импрессионистическая шкала (например, по обычной пятибалльной шкале) по всем объектам тестирования.

В таблице 26 приведены возможные задания для проверки навыков и умений изучающего чтения/аудирования и показатели для оценки качества письма тестируемого.

Прагматические тесты – частичный диктант и копитекст («визуальный диктант») будут рассмотрены в следующем разделе. Комбинированными они являются тогда, когда по ним проводится оценка обученности тестируемых в двух видах речевой деятельности.

8. ПРАГМАТИЧЕСКИЕ ТЕСТЫ

8.1. ОБЩАЯ ХАРАКТЕРИСТИКА

Прагматическими называются такие тесты, с помощью которых определение коммуникативной компетенции (уровня владения ИЯ) производится на основе точного или контекстуально приемлемого восстановления/репродукции лингвистических элементов (букв, слов, предложений) текста, предъявленного визуально или аудитивно, в котором определенным образом уменьшена избыточность информации, причем указанное восстановление/репродукция осуществляется с помощью кратковременной памяти и механизма вероятностного прогнозирования указанных лингвистических элементов текста.

Таблица 26

Пример бланка-схемы оценки результатов комбинированных речевых тестов

Типы заданий (см. табл. 21, все задания со СКО)	Оценки по выбранным объектам тестирования в тексте чтения или аудирования (по табл. 20 или 25)	Критерии оценки письма
	1 12	13 20
1. Формулировка идеи, тем, фактов и т.п.		
2. Вопрос по содержанию		
3. Утверждение (интерпретация с завершением)		
4. Трансформации-парафразы по содержанию		
5. Реферат (свободный или по заданному плану)		
6. Письменное изложение содержания (свободное или по заданному плану)		
7. Прагматические тесты (частичный диктант и копитест)		

Примечание: Критериями оценки письма могут быть: 13 – правописание, 14 – пунктуация, 15 – лексика, 16 – грамматика, 17 – примерная доля правильно переданных фактов, 18 – полнота и правильность передачи основных идей, мыслей и т.п., 19 – оценка за содержание, 20 – оценка за организацию изложения и др.

Отличительными признаками прагматических тестов являются то, что: 1) в них уменьшена избыточность информации; 2) они требуют использования вероятностного прогнозирования как на уровне микроконтекста (предложения, синтагмы), так и на уровне мини- и макроконтекста (СФЕ и текста); 3) они являются интегративными, поскольку: а) требуют понимания и мыслительных операций как на уровне микроконтекста, так и на уровне макроконтекста, б) будучи тестами понимания, они проверяют также лексико-грамматические и фонологически-графологические навыки, память и ряд других способностей, благодаря

чему результаты прагматических тестов имеют весьма хорошую корреляцию с суммарными результатами комплексных тестов, т.е. являются тестами общего уровня владения ИЯ.

Прагматическими тестами называются тесты типа клоуз (ТТК) диктант и некоторые другие (см. табл. 27). Такое название дал этим тестам Дж. Оллер-мл. [136], считая, что в этих тестах используется механизм (грамматика) вероятностного прогнозирования (expectancy grammar). Б. Спольски [153] назвал эти тесты тестами уменьшенной избыточности информации (reduced redundancy test).

Приведенные в таблице 27 разновидности прагматических тестов измеряют в основном рецептивные и репродуктивные навыки и умения. Только в случае ТТК со СКО требуются также некоторые продуктивные навыки. Объекты тестирования в прагматических тестах выражены нечетко, что затрудняет их использование в диагностических целях.

Прагматические тесты имеют довольно высокую степень коммуникативности благодаря достаточно аутентичной ситуации и значительной доли макроконтекстных заданий. В реальной жизни иногда также нужно воспринимать сообщения на фоне шума (объявления на вокзале, переговоры при неисправности телефона и т.п.). Степень коммуникативности можно увеличить путем целенаправленного выбора лингвистических элементов текста, предназначенных для восстановления/репродукции таким образом, чтобы они максимально представляли субкомпоненты и компоненты КК.

Таблица 27

Общие характеристики основных видов прагматических тестов

Название прагматических тестов	Характер текста	Способ уменьшения избыточности информации (вид шума)	Канал восприятия	Действия тестируемого
1. Диктант на фоне шума	предложения, текст	белый шум разной интенсивности	слуховой	написание услышанных частей текста
2. Частичный диктант со СКО или по ТМВ	несколько текстов	пропуск слов, фраз, предложений	слуховой/зрительный	написание в пропуски услышанного
3. Диктант со СКО или по ТМВ	короткий текст – 100 слов	паузы	слуховой	написание услышанных частей текста

Название прагматических тестов	Характер текста	Способ уменьшения избыточности информации (вид шума)	Канал восприятия	Действия тестируемого
4. Диктант-сочинение	длинный текст	паузы	слуховой	письменное изложение услышанного
5. Копитест со СКО («визуальный диктант»)	текст	паузы	зрительный	написание увиденных частей текста
6. Аудиоклоуз-тест со СКО или по ТМВ	текст	пропуски букв, слов, словосочетаний	слуховой	написание или выбор недостающих слов текста
7. Тест типа клоуз со СКО или по ТМВ	текст – 350 слов	пропуски слов	зрительный	написание или выбор недостающих слов текста
8. Б-клоуз-тест со СКО и гибкий Б-тест	текст – 300 слов	пропуски слов, заданы первая и последняя (или только первая) буквы	зрительный	написание недостающей части слова
9. С-клоуз-тест со СКО	короткий текст – 200 слов	пропуски слов, первая половина слов задана	зрительный	написание недостающей части слова
10. Тест побуквенного восстановления текста	короткий текст	неполный текст	зрительный	угадывание слов текста по буквам
11. Частичный или полный перевод без словаря со СКО или по ТМВ	текст	незнакомые слова, несоответствия	зрительный	вероятностное прогнозирование незнакомых слов и перевод
12. Тест редактирования (cloze-elide test)	текст – 300 слов	лишние, неподходящие для данного контекста слова	зрительный	нахождение лишних слов

8.2. ДИКТАНТ

Диктант как прием обучения и контроля навыков и умений аудирования и правописания известен со времен Древнего Рима. Затем он пережил 2–3 периода процветания и с 70-х годов снова используется как интегративный тест аудирования. Ч. Стансфилд цитирует около 40 работ по диктанту [154]. Р. Ладо [127] отрицал значение диктанта для тестирования, считая, что в нем все слова и формы заданы, а сам диктант чересчур сложен и неэффективен. Однако Дж. Оллер-мл. и другие [135; 138; 109] показали, что с помощью диктанта можно надежно измерить понимание при аудировании и уровень общего владения иностранным языком. Диктант включен в состав ряда стандартизованных тестов, а в тест СЛЕП [155] даже в виде теста по ТМВ.

В ходе диктанта тестируемым трижды предъявляют текст в нормальном темпе громкого чтения. Во время второго предъявления текста в определенных местах делаются паузы, в течение которых тестируемые записывают услышанные сегменты (части текста между двумя паузами). Во время третьего предъявления текста также делаются паузы примерно в два раза короче для проверки записи и исправления ошибок. Сегментами служат логически завершенные отрезки текста. Количество слов в сегментах во многом определяет трудность диктанта для данного контингента. Обычно это 7–12 слов, но Г. Цико [102] рекомендует постепенно увеличивать длину сегментов с 3–4 слов до 20–22 слов.

Длительность пауз в секундах определяется делением длины сегментов в печатных знаках на два при втором предъявлении и на пять при третьем предъявлении текста. Чтобы достичь такой длительности, тестирующий может дважды побуквенно прочесть про себя данный сегмент, а затем продолжать громкое чтение следующего сегмента текста.

Оценку диктанта можно производить несколькими способами: 1) учитывая количество правильно написанных слов (один балл за одно слово); 2) учитывая количество безошибочных сегментов при их увеличивающейся длине [102]; 3) принимая во внимание количество лексико-грамматических и семантических ошибок, а также пропусков (но не ошибок правописания). Г. Цико отмечает, что подсчет результатов по правильным сегментам ускоряет процесс проверки диктантов в 3–4 раза. Корреляция между результатами проверок диктантов нескольких рейтеров обычно более 0,90 [146; 85], что предполагает хорошо продуманную процедуру оценивания диктантов.

Диктант на фоне шума и диктант-сочинение (комбинированный тест – после прослушивания текста тестируемые пишут изложение) мало распространены, и об их качествах мало известно. Более перспективным кажется частичный диктант [61], сочетающий диктант и ТТК. В напечатанном тексте опущены предложения или их части, которые нужно восстановить, опираясь на однократное слуховое восприятие текста. Пропуски в тексте лучше поместить ближе к концу предложений. Частичный диктант лучше подготовить по результатам опробова-

ния диктанта, которое позволяет выявить приемлемые пропуски с учетом их трудности и коммуникативного потенциала (количества и характера требуемых компонентов КК).

По сравнению с тестом по аудированию диктант имеет следующие преимущества: бо́льшую степень непосредственности измерения (нет вопросов, просьб типа «сформулируйте основную мысль...» и т.п.), бо́льшую степень коммуникативности (более аутентичная ситуация), проще и быстрее можно подготовить и провести диктант. К недостаткам диктанта относятся трудоемкость, субъективность и проблемность проверки и оценки, сложность интерпретации и несопоставимость результатов разных диктантов, более сложные статистические критерии, поскольку нет одноразмерности результатов теста и некоторые «задания» зависят от успешности решения предшествующих «заданий», особенно в случае частичного диктанта. Последний позволяет использовать более длинные тексты, что увеличивает возможности выбора текста и степень коммуникативности теста. Отбор текста производится аналогично отбору текста для речевых тестов.

8.3. АУДИОКЛОУЗ-ТЕСТЫ

Аудиоклоуз-тестами (АКТ) ниже называются аудитивные тесты восстановления (тесты типа клоуз), в которых в аудиотексте пропущены слова, которые тестируемые должны восстановить во время пауз второго чтения аудиотекста [42; 160; 43; 33]. Места пропусков указываются и во время первого чтения текста. Х. Темплетон [160] предлагает использовать в АКТ шаг-интервал между пропусками в 15 слов и дает только 4 секунды на заполнение пропуска. В экспериментах ТПИ время пауз пришлось увеличивать до 7–8 секунд, чтобы тестируемые успевали записать СКО, и до 12 секунд для заданий по ТМВ.

В ТПИ проведены пилотажные эксперименты с АКТ для выяснения роли трудности текста (по 3 легких и трудных текста), шага между пропусками и типа ответов (СКО или ТМВ). Результаты опытов в таблице 28 показывают, что составленные АКТ оказались легче оптимального (тексты для VIII класса). Трудность текста (по читабельности и проценту НС), шаг и тип ответа оказывают примерно одинаковое не очень сильное влияние на результат АКТ. Оптимальный шаг: 10–14 слов, т.е. в среднем один пропуск в одном предложении. В случае АКТ со СКО трудно исключить «сотрудничество» сидящих рядом тестируемых, эта же проблема возникает и в случае диктанта.

Сравнительно низкие коэффициенты корреляции в таблице 28 связаны с малым количеством заданий в АКТ (до 24 при шаге в 8 слов), с их первым опробованием и с большей ролью памяти в АКТ. В АКТ со СКО больше зависимость от логической (запоминание текстов), ассоциативной и кратковременной вербальной памяти, а в АКТ по ТМВ больше роль механической, логической, кратковременной и оперативной (запоминание цифр) памяти.

Таблица 28

Результаты опробования аудиоклоуз-тестов (группы по 50 первокурсников ТПИ)

Параметры	Процент правильных ответов при шаге между пропусками:		
	8 слов	16 слов	24 слова
Легкий текст, СКО	66	77	81
ТМВ	85	89	–
Трудный текст, СКО	54	67	–
ТМВ	65	86	91
Коэффициент корреляции с результатом теста ИТА-83, АКТ со СКО	0,57	0,61	0,60
по ТМВ	0,46	0,60	0,57

В литературе [121; 90] приведены данные о сравнительно высокой корреляции (0,70–0,80) между результатами тестов устной речи и АКТ со СКО на основе текстов диалогической речи (примеры приведены). В обоих статьях указывается, что нет существенной разницы при подсчете результатов по точным соответствиям (т.е. по словам оригинала) или по контекстуально подходящим СКО. Это установлено и в случае визуальных ТТК.

8.4. РАЗНОВИДНОСТИ ВИЗУАЛЬНЫХ ТЕСТОВ ТИПА КЛОУЗ

Общие положения. В 1897 г. немецкий психолог Х. Эббингхаус предложил для определения умственного развития детей тест, названный несколько позже тестом завершения. В этом тесте в среднем каждое четвертое слово текста было заданием-пропуском, которое нужно было восстановить [115]. Позже, после первой мировой войны, методисты стали использовать пропуски в изолированных предложениях и в меньшей степени для контроля.

В 1953 г. В. Тейлор [159] предложил использовать так называемые клоуз-тесты для определения читабельности текстов, а в 1956 г. и для установления навыков и умений чтения на родном языке. Термин «клоуз-тест» он образовал от английского слова closure *закрытие, смыкание.* Этот термин использовался в начале XX века в гештальт-психологии для установления способности увидеть в изображении предмета с недостающими деталями очертания целостного предмета. Отсюда и переход в методику ИИЯ термина cloze procedure *восстановление недостающих элементов текста.*

В русском языке есть несколько эквивалентов термина «клоуз-тест», например, подстановочный тест, тест (методика) дополнения [62], тест восстановления [71]. В связи с довольно большим числом модификаций (см. табл. 29) клоуз-тестов ниже используются термины «тест типа клоуз» (ТТК) или «клоуз-тест». В преподавании ИЯ ТТК используются в основном для определения общего уровня владения ИЯ, а иногда также в качестве теста по чтению. Иногда получаются низкие коэффициенты корреляции между результатами ТТК и других видов тестов (в том числе по чтению). По-видимому, это происходит из-за значительной зависимости качеств ТТК от характеристик текста, от неполного использования принципов составления ТТК и от того, что ТТК требует от тестируемых различных навыков и умений, субкомпонентов и компонентов КК, а также ряда способностей и различных видов памяти.

Таблица 29

Признаки различных модификаций ТТК

Название ТТК	Тип ответа	Характер шага пропусков	Средний шаг в словах	Тип подсчета (проверки результатов)	Минимальное необходимое количество пропусков
1. Классический	СКО и ТМВ	фиксированный	5, 6, 7	ТО КПС	50
2. Целенаправленная выборка	СКО и ТМВ	нефиксированный	6–10	ТО КПС	40
3. Коммуникативно-ориентированная выборка	СКО	**нефиксированный**	7–12	ТО КПС	35
4. С-тест	СКО	**фиксированный**	2	ТО	100
5. Б-тест	СКО	**нефиксированный**	3–6	ТО	70
6. Комбинированный (гибкий) (п.п. 3, 4, 5)	СКО	**нефиксированный**	4–8	ТО КПС	60

Название ТТК	Тип ответа	Характер шага пропусков	Средний шаг в словах	Тип подсчета (проверки результатов)	Минимальное необходимое количество пропусков
7. Тест редактирования (cloze-elide test)	СКО и ТМВ	нефиксированный	5–10	–	50

Задания-пропуски составляются в случае классического ТТК механическим извлечением из текста каждого *n*-го, например, каждого 5, 6, 7-го и т.д. слова, и проверка результатов ТТК по СКО на начальном и среднем уровне подготовки по ИЯ тестируемых осуществляется по словам, точно совпадающим со словами оригинала (ТС), а на более высоком уровне подготовки по контекстуально подходящим словам (КПС). Экспериментально доказано, что результаты ТТК, проверенные по методике ТС и КПС, имеют очень высокую корреляцию, т.е. взаимозаменяемы.

Недостатки классических ТТК. Классический ТТК, однако, имеет ряд больших недостатков:

1) Пропуски попадают в 10–20 % случаев на слишком трудные слова (цифры, имена собственные и другие), которые не могут угадать даже носители языка, и поэтому очень редко получается стопроцентный результат выполнения ТТК.

2) Интервал между пропусками (шаг) и точка (слово) начала отсчета пропусков могут влиять на структуру частей речи в пропусках таким образом, что не образуется их случайная выборка, у которой было бы соотношение З/С, близкое к соотношению З/С во всем тексте и в языке.

Перед отбором текстов в тест АТУКЛОУЗ-86 был проведен анализ на ЭВМ 24 текстов научно-популярной и научно-технической прозы. Варьировались шаг (2–8 слов) и все возможные 35 точек начала отсчета пропусков. В среднем в 40 % случаев отклонение соотношения знаменательных и служебных слов (З/С) вышло за пределы ± 5 % текстового соотношения З/С (обычно это было между примерно 65 : 35 и 55 : 45). Поскольку экспериментально установлено, что части речи имеют разную трудность в пропусках ТТК, то для получения ТТК с желаемыми параметрами (трудность, ДС и др.) необходимо частично или полностью отказаться от строго фиксированного шага пропусков.

3) Третьим недостатком ТТК является их большая общая трудность. Результат 50 % правильных ответов на вузовском уровне подготовки

при средней трудности текста следует считать достаточно хорошим результатом. Однако нами экспериментально установлено, что если трудность ТТК ниже 40 %, то средний тестируемый будет работать только на уровне лексико-грамматических навыков в пределах нескольких слов до и после пропусков; при этом не используются процессы понимания и механизмы вероятностного прогнозирования, а появляется некая способность использовать локальную подсказку и локальную избыточность информации. При такой работе заметно ухудшается интерес и мотивация тестируемых.

ТТК становится легче на 20–25 % (см. рис. 14) в случае использования заданий-пропусков по ТМВ. Обычно считается, что это уже не ТТК. Наши эксперименты с одинаковыми текстами и контингентами указывают на то, что ТТК по ТМВ обладает всеми качествами ТТК, причем если он имеет трудность ниже 40 %, то появляется вышеуказанное явление «сверхтрудности». В факторном анализе результатов ТТК тогда генеральным фактором становится вместо фактора общего уровня владения ИЯ некоторый клоуз-фактор на основе умений анализа близлежащих к пропуску слов и словосочетаний. Корреляция между ТТК со СКО и по ТМВ составляла в экспериментах ТПИ [32] в среднем 0,60 (10 измерений с текстами; сопоставлялись пары с одинаковыми контингентами, одинаковым фиксированным шагом в 6 слов).

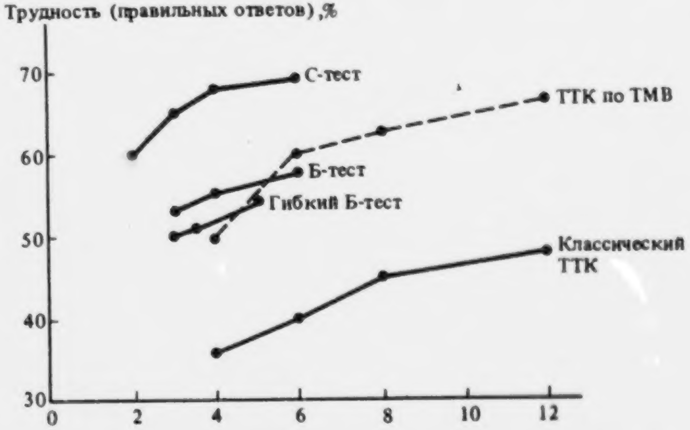

Рис. 14. Сопоставительная трудность разных типов ТТК (усредненные данные 4 текстов и 6 контингентов по 50 студентов, уравненных по результатам теста ИТА-83).

С-тест и Б-тест. Ряд недостатков классических ТТК позволяет устранить С-тест, который предложили в 1981 г. западногерманские ученые У. Раатц и К. Клайн-Брелей [143]. В этом тесте в каждом втором слове

текста пропущена вторая половина слова. Если в слове нечетное число букв, то пропускается бóльшая часть букв. Если слово состоит из одной буквы, то пропускается половина следующего слова (два задания подряд). При такой процедуре обеспечена представительная выборка всех лингвистических элементов на микро- и макроуровнях, нет проблем с шагом, началом отсчета пропусков, подсчетом результатов – только ТС, который является объективным. Учитываются только точно восстановленные слова. Минимальное количество заданий – 100.

Модификацией С-теста является предложенный автором Б-тест (от слова «буква») [125], при котором шаг увеличен до 3–4 и в каждом слове-пропуске указана первая буква. Преимуществом Б-теста является меньшая подсказка с сохранением в большинстве случаев объективности подсчета результатов, а также возможность использования несколько более длинных текстов. Кроме того, при шаге в 3–4 слова имеется больше свободы выбора слова-пропуска, поскольку и в С-тесте иногда желательно избегать слишком легких слов и слов с именами собственными, датами, цифрами, названиями и т.п. При этом сохраняется, с одной стороны, высокая экономичность, а с другой стороны, несколько уменьшено количество заданий.

В ТПИ проведена сопоставительная валидация всех модификаций ТТК [31; 125]. Основным результатом этого исследования является вывод о том, что по общему характеру С-тест и Б-тест не отличаются от классического ТТК.

Как видно из данных рис. 14, при прочих равных условиях труднее всего классический ТТК со СКО и легче всего С-тест. Коэффициенты корреляции с результатами теста ИТА-83 имеют обратный порядок: выше у классического ТТК (0,75–0,85) и ниже у С-теста (0,6–0,7). При увеличении шага эта корреляция также увеличивается у С-теста (шаг 2 видимо мешает более сильной части контингента), а у Б-теста и классического ТТК немного падает. Надежность у всех модификаций порядка 0,80–0,85. Если текст очень трудный, то трудность С-теста также может оказаться сверхтрудной.

Целенаправленная выборка пропусков по принципу, например, включения в тест только знаменательных слов или интуитивный выбор пропусков около мест по фиксированному шагу дает ТТК, который по сравнению с классическим ТТК обычно несколько легче, валиднее и надежнее, поскольку в нем меньше слишком легких и слишком трудных заданий-пропусков с низкой дифференцирующей способностью.

Более целесообразным представляется использование принципа контекстуальности. Согласно этому принципу для пропуска отбирают слова из II и III группы по следующей классификации: восстановление пропущенного слова зависит в I группе от лексико-грамматических связей между словами в словосочетаниях и фразах (например, артикли, предлоги и т.п.); во II группе – от понимания семантико-синтаксических связей между словосочетаниями и фразами в рамках предложения (прилагательные, наречия и т.п.); в III группе – от понимания смысло-

вых связей за пределами одного предложения (например, местоимения, существительные и т.п.); в IV группе – от понимания более отдаленных или внетекстовых смысловых связей. Л. Бахман [86] приводит теоретическое и экспериментальное обоснование этого принципа целенаправленной выборки пропусков ТТК.

Разновидностью ТТК с целенаправленной выборкой пропусков-заданий является ТТК, составленный по принципу коммуникативно-ориентированного выбора локализации заданий ТТК. По этому принципу заданиями ТТК будут слова из вышеуказанных II и III групп, для восстановления которых требуется возможно большее количество разных компонентов КК. Местами заданий ТТК следует также выбирать слова, имеющие лексический повтор, синоним, антоним и т.п. когезивные средства. Установлено, что трудность таких заданий во всех типах ТТК зависит от расстояния (в словах) между повторяющимися (синонимичными) в тексте ТТК словами, одно из которых – задание (коэффициент корреляции 0,70–0,80) [125].

Гибким Б-тестом, или комбинированным ТТК, называется такой ТТК, в котором в каждом задании предложена помощь-подсказка в виде нескольких букв в начале слова, исходя из следующих принципов: 1) букв в начале слова должно быть как можно меньше, ровно столько, чтобы обеспечить однозначность и объективность проверки ответов (нет других контекстуально подходящих ответов с предложенным началом слова); 2) верхним пределом количества неизвлеченных букв (КНБ) в слове является половина слова, как в С-тесте; 3) для I группы контекстуальность КНБ = 0, т.е. задание имеет вид задания классического ТТК; для II и III групп КНБ = 1 – 2 в зависимости от первого принципа и для слов IV группы КНБ = $\frac{n}{2}$ при четном числе букв и КНБ = $\frac{n+1}{2}$ при нечетном числе букв в слове, где n – число букв в слове, как в заданиях С-тестов. Чтобы обеспечить оптимальную трудность заданий (около 60 % правильных ответов), можно после предтеста изменить КНБ; одна буква изменяет трудность задания на 10–15 % по данным рис. 14.

Хорошо зарекомендовал себя редакторский тест, в тексте которого имеются лишние слова, которые нужно найти (выписать). Такой тест имеется в составе стандартизованного теста ТОЭФЛ.

В двух исследованиях [121; 90] показаны возможности оценки навыков устной речи с помощью «устного» ТТК, построенного на базе текста диалогической речи. А. Хьюз [121] использовал шаг в 7, а Д. Браун [156] – в 14 слов. Коэффициенты корреляции с тестами устной речи 0,7 и 0,8 соответственно указывают на необходимость дальнейшего исследования возможностей таких тестов.

Составление тестов типа клоуз. При составлении ТТК возникают как общие проблемы планирования (соответствие характера ТТК целям обучения, уровню подготовки контингента и другие), которые решаются по общим положениям п. 5, так и специфические проблемы отбора и анализа текста, выбора и оформления пропусков и др.

При отборе текста возникает вопрос о знакомстве с текстом. Желательно использовать незнакомый текст, но проверка показала, что повторное проведение ТТК со СКО через неделю повышает результат ТТК на 7–8 %. После беглого чтения текста, на основе которого составлен ТТК, результат ТТК со СКО растет на 15–20 %, если же между чтением и повторным ТТК перерыв составляет одну неделю, то прирост составляет 10–15 %. Если тестируемые читают текст со словарем перед повторным ТТК, то прирост результатов ТТК со СКО составляет 25–32 % [31]. Естественно, что в выигрыше остаются тестируемые с более развитой памятью. Указанные данные говорят о том, что ТТК со СКО можно использовать для проверки домашнего чтения, а ТТК по ТМВ в этом случае слишком легкий.

Для оценки трудности текстов ТТК можно использовать коэффициент читабельности Флеша, поскольку между результатами ТТК и читабельностью текста ТТК обнаружена корреляция 0,84 (14 текстов), а также оценка доли (%) незнакомых слов, которая имеет корреляцию с трудностью ТТК порядка 0,65 [158]. Окончательное решение об отборе текста можно принять только на основе анализа возможных пропусков с точки зрения их структуры частей речи, незнакомых и трудных слов, степени контекстуальности (группы I–IV) и коммуникативности.

Для анализа структуры частей речи и других аспектов можно составить программу для ЭВМ, которая определяет, например, шаг и начальную точку для получения ТТК с нужными показателями. Отсчет пропусков обычно рекомендуется начинать с первого слова второго-третьего предложения текста. Если такой программы для ЭВМ нет, следует решить, какой шаг будет использован, и весь текст переписать в таблицу с колонками для слов (число колонок равно шагу пропусков), чтобы каждая колонка соответствовала пропускам с одной точкой отсчета пропусков. Такое наглядное представление комплектов пропусков-заданий облегчает подсчет и сравнение частей речи (обычно сравниваются количество знаменательных слов и местоимений, незнакомых и трудных слов и т.п.). При целенаправленной выборке пропусков указанная таблица помогает соблюдать выбранный средний шаг.

Незнакомые и трудные слова лучше не пропускать, иначе между пропусками могут быть скрытые пропуски, мешающие контекстуальной догадке. Если основная цель ТТК – проверка чтения, то можно разрешить написать ответы на родном языке.

Составление выборочных ответов для I и II групп заданий (пропусков) аналогично с микроконтекстными заданиями, однако здесь следует учесть, что макроконтекст может сильно повлиять на результат восстановления пропуска. В обратном смысле сложен выбор отвлекающих для заданий (пропусков) III и IV групп, поскольку может обнаруживаться нежелательная подсказка близлежащего микроконтекста (задание нацелено на проверку использования макроконтекстуальных связей).

Задания ТТК оформляются обычно в виде черточек или пунктирной линией, на которых или между которыми указан номер задания. Тестируемые пишут ответы на ответном листе.

9. СОСТАВЛЕНИЕ ТЕСТОВ ПО УСТНОЙ РЕЧИ

С помощью тестов по устной речи (УР) определяется степень развитости навыков и умений говорения, и эти тесты подразделяются на тесты по монологической (МР) и диалогической речи (ДР). Последние относятся к комбинированным тестам, поскольку с их помощью определяется также понимание при слушании – навыки и умения аудирования. Тесты по МР и ДР могут основываться на подготовленной и неподготовленной речи. В последнем случае требуется реакция – ответ на непосредственно заданную ситуацию (стимул).

Для составления тестов по УР можно применить один из трех подходов – традиционный (дискретный), интегративный (контекстуальный) и коммуникативный (см. табл. 30 и 4). Первый подход предполагает косвенное определение навыков и умений УР, т.е. с его помощью определяются лексические, грамматические и просодические навыки. Второй и третий подходы отличаются тем, что при коммуникативном подходе оценка навыков и умений производится по компонентам коммуникативной компетенции (КК).

Таблица 30

Характеристика подходов к составлению тестов по УР

Характеристики тестов	Подходы		
	дискретный	интегративный	коммуникативный
1. Степень коммуникативности	низкая	средняя	высокая
2. Степень непосредственности	низкая, косвенная	средняя, полукосвенная	высокая, непосредственная
3. Степень подготовленности (наличие опор)	высокая	средняя	низкая
4. Показатели для оценки	дискретный	комбинированный	коммуникативный
5. Характер оценки результатов	полуобъективный или объективный	полуобъективный или субъективный	субъективный

В тестах успешности обучения или общего уровня подготовки обычно имеются объекты тестирования как по дискретному (просодические и лексико-грамматические навыки), так и по интегративному подходу. Последние приведены в таблице 31 для теста по МР:

Таблица 31

Объекты тестирования и виды опор в тестах по монологической речи

Объекты тестирования (навыки и умения)	Виды опор			
	прочитанный или прослушанный текст	вопрос, инструкция или план	статические или динамические средства наглядности	физические действия (демонстрация, рисование и т.п.)
1. Ситуативное высказывание	+ +	+	+ +	+ +
2. Высказывание – развернутый ответ на вопрос	+	+ +	–	–
3. Высказывание – описание (фактов, идей и т.п.)	+ +	–	+ +	+ +
4. Высказывание – рассуждение (оценка)	+ +	+ +	+ +	+
5. Передача главной мысли (основного содержания)	+ +	–	+ +	+ +
6. Сообщение по теме	+ +	+ +	+	–
7. Сообщение по развернутому плану	+ +	+ +	+	+
8. Устный реферат или аннотация	+ +	+	+	–
9. Пересказ содержания	+ +	+ +	+ +	+

Примечание: «+ +» – хорошее сочетание; «+» – приемлемое, возможное сочетание; «–» – невозможное или нежелательное сочетание.

К объектам тестирования по диалогической речи относятся: 1) умения, связанные с пониманием реплик собеседника (см. п. 7.2); 2) умения формирования и формулирования кратких или полных ответов на общий, специальный или альтернативный вопросы, которые относятся к определенной ситуации; 3) умение постановки общих, специальных или альтернативных вопросов по определенной ситуации; 4) умение вести (поддержать) беседу на заданную тему, по данной ситуации, используя для этого предвосхищение речевого поведения собеседника и т.п.

В таблице 32 приведены объекты тестирования для коммуникативных тестов по устной речи. Из данных этой таблицы следует, что использованные за 1972–1985 гг. объекты тестирования отражают только часть показателей и субкомпонентов КК. Это связано с неразработанностью этих тестов и с тем, что в один тест невозможно включить весь арсенал показателей и субкомпонентов КК. Ограничения ставят как временные возможности проведения теста (до 30 мин. на тестируемого), так и возможности рейтеров – даже 2 рейтера не смогут оценить более 8–10 показателей. Обычно один тестолог ведет интервью, другой оценивает, или же два рейтера прослушивают 1–2 раза ленты с записью ответов тестируемых.

Таблица 32

Объекты тестирования в зарубежных тестах-интервью

Объекты тестирования	Источник литературы (пор. №)								
	144	150	141	142	89	113	95	164	145
Лингвистическая компетенция									
1. грамматика	+	+	+		+	+		+	+
2. лексика	+	+			+			+	+
3. беглость	+	+	+			+	+	+	+
4. правильность (точность лексико-грамм. форм)				+	+		+		+
5. произношение	+		+					+	
6. понятность			+	+		+			
7. понимание		+		+		+			
Социолингвистическая компетенция						+			
1. вариативность (регистр)					+				
2. уместность							+		
3. естественность речи (nativeness)	+				+				

Объекты тестирования	Источник литературы (пор. №)								
	143	150	141	142	89	113	95	164	145
Дискурсивная компетенция	+			+	+				
1. связность	+								
2. когезивная организация					+				
3. повторяемость для связности							+		
Стратегическая компетенция						+			
1. сложность тем							+		
2. гибкость							+		
3. независимость	+						+		
Страноведческая компетенция					+				
Иллокутивная компетенция – приемлемость							+	+	+

Общими показателями при реализации актов коммуникации являются: 1) их результативность (выполнена ли коммуникативная задача, достигнуто ли коммуникативное намерение); 2) коммуникативная значимость (релевантность) и ценность (новизна, глубина мыслей) речевых актов; 3) экономичность (экономно ли достигнуты цели). Общими показателями для компонентов КК являются: 1) приемлемость формы и содержание (лингвистическая, социолингвистическая и страноведческая компоненты КК) и 2) уместность для данной ситуации общения (стратегическая, дискурсивная и иллокутивная компоненты КК).

Тест-интервью

По данным П. Симмондса [152], интервью и ролевая игра являются наиболее часто используемыми приемами в стандартизованных тестах по говорению (из 12 тестов они используются в 10 и 7 случаях соответственно). В 70-е годы объектами тестирования в этих тестах были лексика, грамматика, произношение, беглость и понимание.

Трудной проблемой для субъективных тестов по устной речи – тестов с субъективной оценкой результатов – является проблема обеспечения стабильности и надежности работы рейтеров. Если рейтер один, то обычно он оценивает результаты интервью с магнитной ленты, ибо вести интервью и сделать записи-оценки весьма трудно. Если рейтеров двое, то обычно один из них ведет интервью и ставит в конце интервью свою оценку по общему впечатлению, а другой рейтер оценивает тестируемого с помощью заранее подготовленной таблицы-сетки (см. табл. 33 и 34).

Таблица 33

Пример таблицы-сетки для оценки теста по устной речи

Показатели, объекты тестирования или субкомпоненты КК	Оценки, баллы, критерий и т.п., например:			
	неудовлетворительно	удовлетворительно	хорошо	отлично
1. Понимание интервьюирующего				
2. Понятливость (коммуникативность) реплик тестируемого				
3. Беглость				
4. Лексико-грамматическая правильность реплик				
5. Другие				

По данным литературы [149; 156; 110; 151], стабильность оценок рейтеров по коэффициенту корреляции колеблется в среднем между 0,80 и 0,90, что можно считать вполне приемлемым.

В таблицу-сетку (табл. 33) могут быть включены другие оценки и критерии, объекты тестирования и показатели могут быть также заменены или дополнены другими объектами тестирования и показателями.

В начале теста-интервью несколько вопросов оставляют без оценки – это простые вопросы для создания атмосферы разрядки и «разминки».

В зависимости от уровня подготовки тестируемых и от целей тестирования тесты-интервью бывают с большей или меньшей степенью управляемости. При меньшей степени управляемости (уровень подготовки выше) тестовые вопросы нацелены на микромонологи, развернутые реплики и т.п. При большей степени управляемости вопросы теста предполагают сравнительно короткие ответы (одно предложение). В этом случае можно для оценки использовать таблицу-сетку, показанную в таблице 34. Вопросы можно градуировать по трудности и составить тест из блоков по 3–5 вопросов каждый. Если тестируемый не смог понятливо ответить на 2–3 вопроса данного блока, то тест можно прекратить, поскольку на еще более трудные вопросы тестируемый вряд ли даст удовлетворительные ответы. Этим можно сэкономить время на проведение теста [32].

Можно использовать более простую шкалу оценок понятности (ком-

муникативности) реплик тестируемого, например, в виде: 0 – непонятный ответ, 1 – понятный, но с заметными ошибками, не мешающими коммуникации, и 2 – ответ с минимальными ошибками. Тогда в таблице 34 вместо показателей можно проставить список тестируемых.

Тест-интервью можно организовать между двумя тестируемыми или в группе из 3–4 тестируемых. Хотя при этом требуется дополнительная подготовка и инструктаж тестируемых, однако экономия времени большая. На каждого тестируемого следует подготовить отдельную таблицу-сетку для оценки результатов теста. К вопросам интервью можно применить обычные процедуры обработки и интерпретации результатов ЛДТ.

Таблица 34

Пример таблицы-сетки для оценки теста-интервью

Показатели, объекты тестирования	Номера вопросов теста-интервью – оценки по ним				Сумма
	1	2	...	15	
1. понимание					
2. понятность реплик тестируемого					
3. беглость					
4. другие					

При подготовке к ролевой игре тестируемым дают ролевые карты с инструкциями, картинами-диаграммами, со списками некоторых реплик, структур, слов и словосочетаний, чтобы лучше управлять ходом проведения теста.

10. ИСПОЛЬЗОВАНИЕ ЭВМ В ЛИНГВОДИДАКТИЧЕСКОМ ТЕСТИРОВАНИИ

Выше было указано, что использование ЭВМ в преподавании ИЯ позволяет сблизить обучение и контроль благодаря быстрой обратной связи, которую ЭВМ может дать в ходе контролирующей программы. При этом используются такие выгодные для обучения и контроля качества ЭВМ, как их быстродействие, большой объем памяти, возможность вербальных и невербальных модификаций на экране ЭВМ: замена, дополнение – преобразование, редактирование и т.п. поданных на экран ЭВМ элементов программы.

При использовании ЭВМ для тестирования желательно соблюдать

следующие основные принципы: 1) максимально возможная контекстуальность заданий теста (мини- и макроконтекстуальные задания); 2) максимальное использование интерактивности – ЭВМ дает дополнительные инструкции, примеры, намеки при затруднениях и ошибках тестируемых; 3) преимущественное использование заданий со СКО, но поскольку ввод ответа часто затруднен из-за объема или правописания, то лучше использовать короткие ответы (слово или словосочетание).

Оценка ответов тестируемых может быть произведена: 1) с приведением правильного ответа; 2) с указанием сразу правильности – неправильности ответа; 3) с указанием места и вида ошибки, например, в случае ошибки в правописании; 4) с приведением комментария, правил, примеров в случае ошибки.

Почти все типы заданий можно использовать в компьютеризированном ЛДТ. Однако проверку навыков и умений чтения и аудирования можно реализовать в основном с помощью заданий по ТМВ или со СКО полуоткрытого типа или с коротким ответом.

Удобно использовать ЭВМ для визуальных прагматических тестов. Например, в копитесте можно на короткое время предъявить на экране текст, а его восстановление может быть осуществлено так, как в частичном диктанте – часть текста (начало предложения, имена собственные, цифры, даты и т.п.) остается на экране, а остальное тестируемый вводит в ЭВМ. В тестах типа клоуз ЭВМ может по ходу теста согласно результатам первых 5–10 заданий перейти на более трудный или более легкий вариант теста (увеличивая шаг между пропусками или увеличивая подсказку – давая первую букву или полслова в пропуске), чтобы избегать сверхтрудности ТТК для данного тестируемого. Такие программы можно назвать адаптирующими.

Поскольку программирование занимает много времени (30–100 часов на один час работы учащегося с ЭВМ), то следует разработать или использовать так называемые авторские программы, в которых оставлены места для тестируемого/обучаемого учебного материала, который преподаватель желает использовать. Основные операции в таких программах уже запрограммированы.

БИБЛИОГРАФИЯ

1. *Анастази А.* Психологическое тестирование: В 2 т / Под ред. К.М. Гуревича и В.И. Лубовского. М., 1982.

2. *Антонова И.И.* Вопросы контроля умений в монологической устной речи учащихся: Дисс. ... канд. пед. наук. М., 1967.

3. *Банкевич Л.В.* Система тестов в американской методике // Иностр. яз. в школе. 1965. № 6. С. 51–61.

4. *Банкевич Л.В.* Тестирование лексики иностранного языка. М., 1981.

5. *Веденяпин Ю.А.* Измерение готовности абитуриентов к обучению в институ-

тах иностранных языков как предпосылка индивидуализации учебного процесса: Дисс. ... канд. пед. наук. М., 1976.

6. *Веденяпин Ю.А.* Методические указания (письмо) № 7 по вопросам программированного обучения иностранным языкам в вузах. МГПИИЯ, 1971 г. Ч. I, II.

7. *Витковская Ж.А.* Методические требования к тестам для проверки уровня владения ознакомительным чтением на иностранном языке: Дисс. ... канд. пед. наук. М., 1975.

8. *Воскерчьян С.И.* Об использовании метода тестов при учете успеваемости школьников // Советская педагогика. 1963. № 10. С. 28–37.

9. *Вырк Э.-М., Коккота В.А.* Временной фактор в языковых тестах // Уч. зап. Тартуского ГУ: Труды по методике преподавания иностранных языков. Тарту, 1978. Вып. 7. С. 35–40.

10. *Гальцев И.Н.* Летние курсы английского языка в Джорджтаунском университете // Обучение иностранным языкам в высшей школе. 1967. № 3/8. С. 242–248.

11. *Гез Н.И.* Формирование коммуникативной компетенции как объект зарубежных методических исследований // Иностр. яз. в школе. 1986. № 2. С. 17–24.

12. *Гильбух Ю.З.* Метод психологических тестов: сущность и знание // Вопр. психологии. 1986. № 2. С. 30–42.

13. *Гохлернер М.М., Рапопорт И.А., Соттер И.* О диагностике способностей к изучению иностранного языка // Проблемы обучения иностранным языкам. Таллинн, 1978. С. 28–43.

14. *Гутман И.И.* Методические указания к тестовой методике контроля текущей успеваемости. Одесса, 1975.

15. *Давыдкина Н.А.* Организация тестовой методики контроля навыков структурного оформления предложений в 5–6 классах средней школы: Дисс. ... канд. пед. наук. Воронеж, 1973.

16. *Довгалевская Н.А.* Системный анализ факторов учебной успешности (на материале обучения иностранным языкам в морском вузе): Дисс. ... канд. пед. наук. Л., 1974.

17. *Зимняя И.А.* Речевая деятельность и речевое поведение в обучении иностранному языку: Сб. научных трудов МГПИИЯ. М., 1984. № 242. С. 3–10.

18. *Иванова Г.В.* Задание тестового типа как один из способов проверки знаний и умений учащихся: Дисс. ... канд. пед. наук. М., 1973.

19. *Иванова Е.Ф.* Организация следящего контроля трудноусваиваемой лексики на I курсе языкового факультета педвуза с помощью тестовой методики (на материале английского языка): Дисс. ... канд. пед. наук. М., 1975.

20. *Ильина Т.А.* Тестовая методика проверки знаний и программированное обучение // Советская педагогика. 1967. № 2. С. 122–125.

21. *Кабардов М.К.* Роль индивидуальных различий в успешности обучения иностранным языкам: Дисс. ... канд. психол. наук. М., 1983.

22. *Калиниченко А.С. и др.* Методические указания по тестовому контролю знаний. Харьков, 1981.

23. *Князева Г.Ю.* Система языковых тестов в США // Уч. зап. / МГПИИЯ. М., 1968. Т. 44. С. 271–278.

24. *Коккота В.А.* Проблемы составления лексических тестов исходного уровня // Вопросы преемственности в преподавании иностранного языка в средней и высшей школе. Свердловск, 1979. С. 45–55.

25. *Коккота В.А.* Исследование качества тестов исходного уровня по английскому языку // Вопросы преемственности в преподавании иностранного языка в средней и высшей школе. Свердловск, 1981. С. 146–159.

26. *Коккота В.А.* Исследование разных видов клоуз-тестов // Уч. зап. / Тарту-

ского ГУ. Вып. 565; Труды по методике преподавания иностранных языков. Тарту, 1981. Вып. 10. С. 40–46.

27. *Коккота В.А.* Исследования языковых способностей первокурсников ТПИ // Уч. зап. / Тартуского ГУ. Вып. 473; Труды по методике преподавания иностранных языков. Тарту, 1978. Вып. 7. С. 35–40.

28. *Коккота В.А.* Методика определения с помощью теста уровня знаний студентов, поступивших в ТПИ // Уч. зап. / Тартуского ГУ. Вып. 345; Труды по методике преподавания иностранных языков. Тарту, 1975. Т. 4. С. 27–36.

29. *Коккота В.А., Вырк Э.-М.* Исследование учебной трудности иноязычной лексики // Система обучения иностранным языкам в неязыковых вузах: Тезисы докладов совещания-семинара зав. кафедрами иностранных языков химико-технологических, горных и других вузов. М., 1978. С. 66–68.

30. *Коккота В.А. и др.* Отчет НИР № 81040157: Исследование исходного уровня подготовки по английскому языку в вузах СССР. ТПИ, 1985.

31. *Коккота В.А. и др.* Отчет НИР № 8104057, инв. № 02840027714: Исследование свойств тестов дополнения. ТПИ, 1983.

32. *Коккота В.А. и др.* Отчет НИР № 929013, инв. № 78080519: Исследование системы показателей для контроля в учебном процессе по иностранному языку. ТПИ, 1980.

33. *Коккота В.А.* Отчет НИР № 8104057, инв. № 02860025З669: Исследование средств контроля аудирования для языковых лабораторий. ТПИ, 1985.

34. *Коккота В.А.* Поиски экономного языкового теста // Теория и практика языковых тестов. Таллинн, 1980. С. 72–80.

35. *Коккота В.А.* Психологический анализ обученности и обучаемости по иностранным языкам // Уч. зап. / Тартуского ГУ. Вып. 539; Труды по методике преподавания иностранных языков. Тарту, 1980. Вып. 9. С. 75–89.

36. *Коккота В.А.* Результаты теста исходного уровня и теста дополнения для I–II курсов в вузах ЭССР // Уч. зап. / Тартуского ГУ. Вып. 617; Обучение разным видам речевой деятельности и проблемы контроля. Тарту, 1982. С. 48–57.

37. *Коккота В.А.* Стандартизация теста исходного уровня по английскому языку в вузах ЭССР // Уч. зап. / Тартуского ГУ. Вып. 505; Труды по методике преподавания иностранных языков. Тарту, 1979. Т. 8. С. 55–65.

38. Контроль навыков и умений в основных видах речевой деятельности по иностранному языку в неязыковых вузах // Сб. научн. трудов / МГПИИЯ. М., 1980. № 153.

39. *Королева Т.А.* О тестировании в интенсивном курсе иностранных языков (Франция) // Обучение иностранным языкам в высшей школе: Методический сборник. М., 1969. 4/9. С. 128–139.

40. *Красюк Н.И.* Методика тестового контроля навыков лексико-грамматического оформления устного иноязычного высказывания (I курс языкового вуза): Дисс. ... канд. пед. наук. М., 1981.

41. *Красюк Н.И.* Пособие для преподавателей иностранных языков по организации тестового контроля. Днепропетровск, 1976.

42. *Кунина Е.А.* Использование тестовой методики для аудирования на начальном этапе обучения в языковом вузе (английский язык): Дисс. ... канд. пед. наук. М., 1980.

43. *Кунина Е.А.* Использование устного теста восстановления для контроля понимания иноязычной речи на слух // Иностр. яз. в высш. школе. М., 1980. Вып. 15. С. 25–32.

44. *Курдюмова И.М.* Методика комплексного контроля в обучении английскому языку в средней школе: Дисс. ... канд. пед. наук. М., 1975.

45. Методические указания к составлению учебника иностранного языка для вузов неязыковых специальностей. М., 1979. С. 43–56.

46. Методические указания по организации контроля в системе обучения иностранным языкам. М., 1981.

47. *Половой А.Н.* Использование тестирования для выявления методической готовности студентов факультетов иностранных языков к педагогической практике предвыпускного курса: Дисс. ... канд. пед. наук. М., 1983.

48. *Мурумете С. и др.* Исследование форм предъявления лексических единиц в тесте исходного уровня // II зональное совещание заведующих кафедрами иностранных вузов БССР, ЭССР, Латв. ССР, Лит. ССР и Калининградской обл. (РСФСР) : Тез. докладов. Тарту, 1977. С. 82–84.

49. *Носенко Э.Л.* Объективные показатели уровня владения устной иноязычной (монологической) речью: Дисс. ... канд. пед. наук. М., 1970.

50. Объекты и формы контроля при обучении иностранным языкам // Сб. научн. тр. / МГПИИЯ. 1983. № 223.

51. Опыт тестирования по иностранным языкам в университетах США / Сост. С.П. Макушева. М., 1971.

52. Опыт тестового контроля по иностранным языкам в школе / Сост. А.Ю. Горчев. М., 1977.

53. *Педанова М.А.* Методическая классификация лексики и ее использование при обучении немецкому языку в техническом вузе: Дисс. ... канд. пед. наук. М., 1969.

54. Проблема контроля при обучении иностранным языкам в вузе (языковые тесты: теория и практика) / Сост. И.А. Патурова. Таганрог, 1972. Вып. 1; 1974. Вып. 2; Саратов, 1974.

55. Проблемы повышения эффективности обучения иностранным языкам / Сост. Р. Сельг и И. Соттер. Таллинн, 1983.

56. *Рабинович Ф.М.* Контроль на уроке иностранного языка // Иностр. яз. в школе. 1987. № 1. С. 10–16.

57. *Рапопорт И.А., Махлах Е.С.* К проблеме изучения интеллектуальных компонентов обучаемости иностранному языку // Проблемы обучения иностранным языкам. Таллинн, 1978. С. 6–27.

58. *Рапопорт И.А.* О валидности и надежности педагогических языковых тестов. Таллинн, 1980. С. 7–17.

59. *Рапопорт И.А., Махлах Е.С.* О соотношении обученности иностранному языку с развитием разных видов памяти у школьников старшего возраста // Проблемы повышения эффективности обучения иностранным языкам. Таллинн, 1983. С. 97–119.

60. *Рапопорт И.А.* Об использовании метода тестов при учете успеваемости школьников: Дисс. ... канд. пед. наук. М., 1971.

61. *Рапопорт И.А.* Прагматические тесты: сущность, специфика, перспективы // Иностр. яз. в школе. 1985. № 2. С. 70–74.

62. *Рапопорт И.А., Сельг Р., Соттер И.* Тесты в обучении иностранным языкам. Таллинн, 1987.

63. *Рапопорт И.А., Соттер И.* Опыт разработки стандартизованного теста // Иностр. яз. в школе. 1977. № 1. С. 27–36.

64. *Рапопорт И.А., Соттер И.* Словарные тесты: методика отбора языкового материала и составление // Проблемы психологической диагностики. Таллинн, 1977. С. 32–49.

65. Рекомендации к организации учебного процесса по иностранным языкам на неязыковых специальностях вузов // Организация контроля. М., 1985. С. 21–26.

66. *Решетов А.Г.* Анализ результатов тестирования в ряде неязыковых вузов // Обучение иностранным языкам в высшей школе. М., 1973. Вып. 79. Ч. II. С. 334–346.

67. *Розенкранц М.В.* Использование тестовой методики при обучении чтению в старших классах средней школы (английский язык): Дисс. ... канд. пед. наук. М., 1978.

68. *Свешникова Л.А.* Стандартизация контроля знаний студентов как средство его оптимизации: Дисс. ... канд. пед. наук. Л., 1977.

69. *Сельг Р.Л.* Оптимизация контроля обученности предметам языкового цикла: Дисс. ... канд. пед. наук. Тарту, 1984.

70. Содержание и методы обучения иностранному языку в вузах неязыковых специальностей (третий этап обучения) // Организация контроля. М., 1986. С. 39–47.

71. *Соомере К.Э.* Использование тестов восстановления в обучении иностранным языкам: Дисс. ... канд. пед. наук. М., 1986.

72. *Суворов С.П.* Педагогические тесты: практика их составления и использования в США // Пути улучшения преподавания иностранных языков. М., 1970. С. 51–117.

73. Тезисы всесоюзной конференции «Проблемы и методы исследования качественных и количественных характеристик знаний, навыков и умений учащихся». М., 1976.

74. Теория и практика языковых тестов / Сост. Р. Сельг и И. Соттер. Таллинн, 1980.

75. *Фельснер И.В.* Методика усвоения лексики для чтения на I курсе неязыкового вуза (английский язык): Дисс. ... канд. пед. наук. М., 1976.

76. *Фоломкина С.К.* Тестирование в обучении иностранному языку // Иностр. яз. в школе. 1986. № 2. С. 16–20.

77. *Цатурова И.А.* Из истории развития тестов в СССР и за рубежом. Таганрог, 1969.

78. *Цатурова И.А.* Функции теста как средства контроля // Проблема контроля при обучении иностранным языкам в вузе. Саратов, 1984. С. 3–7.

79. *Цатурова И.А.* Элементы тестирования в методике обучения в технических вузах: Автореф. дис. ... канд. пед. наук. М., 1971.

80. *Шебеко Н.С.* Контроль в системе обучения иностранным языкам в вузе // Сб. научн. тр. / МГПИИЯ. 1989. № 153. С. 9–20.

81. *Шебеко Н.С.* Контроль лексического материала на английском языке в 5–6 классах средней школы: Дисс. ... канд. пед. наук. М., 1975.

82. *Штульман Э.А. и др.* Тесты, требования к ним и их классификация // Тестирование при обучении иностранным языкам в неязыковом вузе / Под ред. П.А. Баскутиса. Воронеж, 1972. С. 7–63.

83. *Юдис А.Б.* Использование тестов для контроля умений студентов читать научно-техническую литературу: Дисс. ... канд. пед. наук. Киев, 1975.

* * *

84. *Allen J.P.A., Davies A.* [eds.] Testing and experimental methods // Edinburgh Course in Applied Linguistics. London, 1977. Vol. 4.

85. *Arnaud P.* La place de la dictée dans une batterie d'entrée // Les langues modernes. P., 1984. No.4. P. 281-291.

86. *Bachman L.* Performance on Cloze Tests with Fixed-ratio and Rational Deletions // TESOL Quarterly. Wash., 1985. Vol. 19. No. 3. P. 535-556.

87. *Bachman L.* The trait structure of cloze test scores // TESOL Quarterly. Wash., 1982. Vol. 16. No. 1. P. 61–70.

88. *Bachman L., Palmer A.* Fundamental Considerations in the Measurement of Language Abilities: Reading. Mass., 1986.

89. *Bachman L., Palmer A.* The Construct Validation of Some Components of Communicative Proficiency // TESOL Quarterly. Wash., 1982. Vol. 16. No. 4. P. 449-465.

90. *Brown D.* Conversational close tests and conversational ability // English Language Teaching Journal. 1983. Vol. 37. No. 2. 158-161.

91. *Bubenikova L. et al.* Teorie a praxe jazykového testováni. Praha, 1975.

92. *Canale M., Swain M.* Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing // Applied Linguistics. 1980. Vol. 1. P. 1-47.

93. *Carroll J.B.* Make Your Own Tests: A practical guide to writing performance tests. Pergamon Press. 1985.

94. *Carroll J.B. Sapon S.M.* Modern Language Aptitude Test. N.Y., 1958.

95. *Carroll B.J.* Testing Communicative Performance: An Interim Study. London, 1980.

96. *Cattel McKeen.* Mental Tests and Measurements. Mind, 1980.

97. *Clark J.L.D.* [ed.] Direct Testing of Speaking Proficiency: Theory and Application. Princeton, 1978.

98. *Clark J.L.D.* Foreign-Language Testing: Theory and Practice. Philadelphia, 1972.

99. *Clark J.L.D.* Language Testing: Past and Current Status – Directions for the Future // The Modern Language Journal. 1983. Vol. 67. No. 4. P. 431-443.

100. *Clifford R.* Testing Oral Language Proficiency: A Dynamic Model // Practice and Problems in Language Testing 2 / Ed. by Chr. Klein-Braley, D. Stevenson. Frankfurt-am-Main, 1985.

101. *Cronbach L.* Essentials of Psychological Testing. N.Y., 1970.

102. *Cziko G.* The Construction and Analysis of Short Scales of Language Proficiency: classical psychometrics, latent trait and nonparametric approaches // TESOL Quarterly. 1984. Vol. 18. No 4. P. 627-647.

103. *Davies A.* Language Testing // Survey Article in Language Teaching and Linguistics Abstracts. 1978. Vol. 11. Part I. No 3. P. 145–159; Part II. No 4. P. 215-231.

104. *Davies A.* [ed.] Language Testing Symposium. London, 1968.

105. *Dubois P.H.* A test-dominated society: China 115 B.C. – 1905 A.D. // Anastasi A. Testing Problems in Perspective. Wash., 1966. P. 29-36.

106. *Faerch C., Casper G.* Plans and Strategies in Foreign Language Communication // Faerch C., Caspar G. [eds.] Strategies in Interlanguage Communication. London, 1983.

107. *Finocchiaro M., Brumfit Ch.* The Functional-Notional Approach: From Theory to Practice. London, 1983.

108. *Finocchiaro M., Sako S.* Foreign Language Testing: A Practical Approach. N.Y., 1983.

109. *Fouly K., Cziko G.* Determining the reliability, validity and scalability of the graduated dictation test // Language Learning. 1985. Vol. 35. No 4. P. 555-566.

110. *Francis J.* The reliability of two methods of marking oral tests in Modern Language Examinations // British Journal of Language Teaching. 1981. Vol. 19. No 1. P. 15-23.

111. *Gerberich J.R.* Specimen Objective Test Items: A Guide to Achievement Test Construction. N.Y., 1956.

112. *Goodrich H.* Distractor Efficiency in Foreign Language Testing // TESOL Quarterly. 1977. Vol. 11. No. 1. P. 69-78.

113. *Grenier-Henrie A.-M. et al.* La mesure de la compétence de communication orale // Medium. Ottawa, 1982. Vol. 7. No 2. P. 113-133.

114. *Halliday M.A.K.* Language as Social Semiotic. London, 1979.

115. *Harris D.* Some Forerunners of Cloze Procedure //Modern Language Journal. 1985. Vol. 69. No 4. P. 367-376.

116. *Harris D.* Testing English as a Second Language. N.Y., 1969.

117. *Harrison A.* Language Testing Handbook. London, 1983.

118. *Heaton J.* Language Testing. London, 1982.

119. *Heaton J.* Writing English Language Tests. London, 1975.

120. *Henmon V.* Achievement Tests in the Modern Foreign Language. N.Y., 1929.

121. *Hughes A.* Conversational cloze as a measure of oral ability // English Language Teaching Journal. 1981. Vol. 25. No 2. P. 161-168.

122. *Hymes D.* On Communicative Competence // Sociolinguistics / ed. by J.B. Pride, J. Holmes. Harmondsworth, 1972.

123. *Ingram E.* Attainment and Diagnostic Testing // Davies A. [ed.] Language Testing Symposium. London, 1968. P. 70-98.

124. *Klein-Braley C.* A cloze-up on the C-test: a study in the construct validation of authentic tests // Language Testing. 1985. Vol. 3. No. 1. P. 76-105.

125. *Kokkota V.* Letter-deletion procedure: a flexible way of reducing the text redundancy for cloze tests // Summaries of the III Regional Seminar "Theoretical Problems of Language Testing." Tallinn, 1988. P. 51-54.

126. *Komorowska H.* Linguistic and Communicative Competence of School Leavers: A Diagnostic Study // Glottodidactica, 1979. No 12. P. 65-78.
127. *Lado R.* Language Testing. London, 1961.
128. *Lange D., Clifford D.* Testing in Foreign Language, ESL and Bilingual Education, 1966-1979. A selected annotated ERIC bibliography. Arlington, 1980.
129. *Lewis E.G., Massad C.E.* The Teaching of English as a Foreign Language in Ten Countries // IEA UNESCO International Studies in Evaluation. Stockholm, 1975.
130. *Lienert P.* Testaufbau und Testanalyse. Berlin, 1969.
131. *Lück H.* The development of the Duisburg English Language Test for Advanced Students (DELTA) // Grotjahn R., Hopkins A. [eds.] Research in Language Learning and Teaching: Quantitative Linguistics 6. Bochum, 1980. P. 85-126.
132. *Madsen H.* Techniques in testing. N.Y., 1983.
133. *Morrow K.* Communicative Language Testing: revolution or evolution? // Brumfit C., Johnson K. [eds.] Communicative Approach to Language Teaching. Oxford, 1981. P. 143-157.
134. *Munby J.* Communicative Syllabus Design: A sociolinguistic model for defining the content of purpose-specific language programs. Cambridge, 1978.
135. *Oller J.* Dictation as a Device for Testing Foreign Language Proficiency // English Language Teaching Journal. 1971. Vol. 25. P. 254-259.
136. *Oller J.* Language tests at school: a pragmatic approach. London, 1979.
137. *Oller J., Richards J.* [eds.] Focus on the Learner: Pragmatic Perspectives for the Language Teachers. Rowley, 1973.
138. *Oller J., Streiff V.* Dictation – a Test of Grammar Based Expectances // English Language Teaching Journal. 1975. Vol. 29. No 1. P. 25-36.
139. *Omaggio A.* Proficiency-oriented classroom testing. Wash., 1985.
140. *Palmer A.* Testing Communication // IRAL. 1972. Vol. 10. No 1. P. 35-45.
141. *Perkins K. et al.* Evidence for unique articulatory and auditory dimensions in ESL data // Kohonen V. et al. [eds.] AFINLA series. 1985. No 40. P. 89-103.
142. *Politzer R., McGroarty M.* A discrete-point test of communicative competence // IRAL. 1983. Vol. 21. No 3. P. 182-191.
143. *Raatz V.* Investigating dimensionality of language tests – a new solution to an old problem // Kohonen V. et al. [eds.] AFINLA series. 1985. Vol. 40. P. 123-136.
144. *Raatz V., Klein-Braley C.* The C-test – a modification of the cloze procedure // Culhane T. et al. [eds.] Practice and Problems in Language Testing // Occasional papers / Univ. of Essex. 1981. No 26. P. 95-117.
145. *Savignon S.* Communicative Competence: An Experiment in Foreign Language Teaching. Philadelphia, 1972.
146. *Savignon S.* Dictation as a measure of communicative competence in French as a second language // Language Learning. 1982. Vol. 32. No 1. P. 33-51.
147. *Savignon S.* Communicative Competence: Theory and Classroom Practice. Mass., 1983.
148. *Schonell F.J., Schonell F.E.* Diagnostic and attainment testing. Edinburgh – London, 1956.
149. *Schulz R.* Discrete-point Versus Simulated Communication Testing in Foreign Languages // Modern Language Journal. 1977. Vol. 61. No 3. P. 94-101.
150. *Shohamy E.* An alternative model for assessing language proficiency in secondary school // Kohonen V. et al. [eds.] AFINLA series. 1985. No 40. P. 149-157.
151. *Shohamy E.* The stability of oral proficiency evaluations // The First International Conference on frontiers in language proficiency and dominance testing. Carbondale, 1977. P. 133-142.
152. *Simmonds P.* A Survey of English Language Examinations // English Language Teaching Journal. 1985. Vol. 39. No 1. P. 33-42.
153. *Spolsky B.* Language Testing: Art of Science: Main lecture delivered at AILA World Congress 1975 // Spolsky B. [ed.] Approaches to Language Testing. Arlington, 1978.
154. *Stansfield Ch.* A History of Dictation in Foreign Language Teaching and Testing // Modern Language Journal. 1985. Vol. 69. No 2. P. 121-127.

155. *Stansfield Ch.* Reliability and Validity of the Secondary Level English Proficiency (SLEP) Test // System. Vol. 11. No 1. P. 56-69.

156. *Strecker B.* Verfahren zur Messung freier Sprechfertigkeit in der Fremdsprache // Die Neueren Sprachen. 1980. V. 80. No 1. S. 125-142.

157. Summaries of the I Regional Seminar: Testing as a means of increasing the efficiency of foreign language teaching / Ed. by A. Pihlak, V. Kokkota. Tallinn, 1984.

158. Summaries of the II Regional Seminar: Testing in foreign language teaching / Ed. by V. Kokkota. TPI, 1986.

159. *Taylor W.* Cloze Procedure: A New Tool for Measuring Readability // Journalism Quarterly. 1953. Vol. 30. No 3. P. 415-433.

160. *Templeton H.* A new technique for measuring listening comprehension // English Language Teaching Journal. 1977. Vol. 31. No 4. P. 292-299.

161. Test of English as a Foreign Language (TOEFL) Interpretive Information. Princeton, 1931.

162. *Thorndike R.* The Measurement of Intelligence. N.Y., 1926.

163. *Upshur J.* Michigan Test of English Language Proficiency, Form A, B. Univ. of Michigan, 1961.

164. *Valdman A., Moody M.* Testing communicative ability //French Review. 1985. Vol. 39. No 1. P. 552-561.

165. *Valette R.* Modern Language Testing. N.Y., 1977.

166. *Valette R.* Modern Language Testing: A Handbook. N.Y., 1967.

167. *Van Ek, Alexander L.G.* Threshold Level English. Oxford, 1980.

168. *Whiteson V.* Foreign Language Testing: A Current View // English Language Teaching Journal. 1981. Vol. 25. No 3. P. 345-352.

169. *Widdowson H.G.* Teaching Language as Communication. Oxford, 1978.

170. *Wilkins D.A.* Notional syllabuses. London, 1976.

СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ

АКТ – аудиоклоуз-тест

АТУ – английский тест успешности, стандартизованный в неязыковых вузах тест для студентов II курса

АЯ – английский язык

Б – тест, буквенный тест, разновидность ТТК

В – (чтение) вслух

ГФ – грамматическая форма

ДВ – двойной выбор (ответов)

ДК – дискурсивная компетенция

ДР – диалогическая речь

ДС – дифференцирующая способность (заданий)

З/С – (отношение количеств) знаменательных и служебных слов (в тексте)

ИЛ – иллокутивная компетенция

ИИ – (коэффициент) информационной насыщенности

ИТ – индекс трудности (заданий)

ИТА – исходный тест английский, стандартизованный в неязыковых вузах тест

ИЯ – иностранный язык

КК – коммуникативная компетенция

КНБ – количество неизвлеченных букв

КПС – контекстуально подходящие слова (в пропуск клоуз-тест)

ЛГ – лексические и грамматические (субкомпоненты)

ЛДТ – лингводидактический тест

ЛЕ – лексическая единица

ЛК – лингвистическая компетенция

ЛСТ – локально стандартизованный тест

МР – монологическая речь

НС – незнакомые слова

ОТ – окончательный тест

ПИЯ – преподавание иностранных языков

ПТ – преподавательский тест

РоЯ – родной язык

РуЯ – русский язык

С – (чтение) про себя

СК – стратегическая компетенция

СКО – свободно конструируемый ответ

СЛК – социолингвистическая компетенция

СТ – стандартизованный тест

СФЕ – сверхфразовое единство

ТМВ – техника множественного выбора (тип заданий)

ТПИ – Таллинский политехнический институт

ТС – точно совпадающие (слова)

ТТК – тест типа клоуз

УР – устная речь

ФЗ – функциональное значение

ЭЯ – эстонский язык

СОДЕРЖАНИЕ

1. Предисловие 3

2. Место тестирования в преподавании иностранных языков 5

2.1. Основные понятия в области лингводидактического тестирования . . . 5
2.2. Роль тестирования в преподавании иностранных языков 9
2.3. Функции лингводидактических тестов 10

3. Теоретические проблемы лингводидактических тестов 13

3.1. Общая характеристика лингводидактических тестов 13
3.1.1. Сравнение ЛДТ с психологическими и педагогическими тестами 13
3.1.2. Преимущества и недостатки лингводидактических тестов 14
3.2. Лингвистическая и коммуникативная компетенция 17
3.3. Теоретические проблемы измерения коммуникативной компетенции 21
3.4. Классификация лингводидактических тестов 26
3.4.1. Языковые и речевые тесты 26
3.4.2. Стандартизованные и нестандартизованные тесты 31
3.4.3. Классификация тестов по целям тестирования 31
3.5. Показатели качества лингводидактических тестов 35
3.6. Экономичность лингводидактических тестов 41

4. Краткая история развития лингводидактического тестирования 49

4.1. Развитие тестирования за рубежом 49
4.2. Развитие лингводидактического тестирования в СССР 54

5. Общая методика разработки лингводидактических тестов 57

5.1. Планирование лингводидактического тестирования 57
5.2. Составление заданий лингводидактических тестов 60
5.3. Обработка и интерпретация результатов теста 67
5.4. Составление окончательного теста (перепланирование предтеста) 72

6. Составление тестов по лексике и грамматике 73

6.1. Общие положения 73
6.2. Составление тестов по лексике 75

6.3. Составление тестов по грамматике 82

7. Составление тестов по чтению и аудированию 85

7.1. Общие положения 85
7.2. Составление тестов по чтению 88
7.3. Составление тестов по аудированию 94
7.4. Комбинированные речевые тесты 98

8. Прагматические тесты 98

8.1. Общая характеристика 98
8.2. Диктант 102
8.3. Аудиоклоуз-тесты 103
8.4. Разновидности визуальных тестов типа клоуз 104

9. Составление тестов по устной речи 111

10. Использование ЭВМ в лингводидактическом тестировании 116

Библиография 117
Список использованных сокращений 125

Учебное издание

КОККОТА Вальмар Александрович

ЛИНГВОДИДАКТИЧЕСКОЕ ТЕСТИРОВАНИЕ

Зав. редакцией *И.Э. Волкова*
Редактор *Л.Н. Белая*
Младший редактор *Е.П. Политова*
Художественный редактор *В.И. Пономаренко*
Технический редактор *В.М. Романова*
Старший корректор *З.Ф. Юрескул*
Оператор *Г.А. Шестакова*

ИБ № 7950

Изд. № А–90. Сдано в набор 07.02.89. Подп. в печать 30.03.89. Формат 60×88/16. Бум. офс. № 2. Гарнитура Пресс-Роман. Печать офсетная. Объем 7,84 усл. печ. л. 8,09 усл. кр.-отт. 7,71 уч.-изд. л. Тираж 7000 экз. Зак. №2275. Цена 30 коп.

101430, Москва, ГСП-4, Неглинная ул., д. 29/14.
Издательство «Высшая школа».
Набрано на наборно-пишущих машинах издательства.

Отпечатано в Московской типографии № 4 «Союзполиграфпрома» при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 129041, Москва, Б. Переяславская ул., 46.